

# Adaptive evolution of molecular phenotypes

Torsten Held<sup>1</sup>, Armita Nourmohammad<sup>1,2</sup>, Michael Lässig<sup>1</sup>

<sup>1</sup> Institut für Theoretische Physik, Universität zu Köln, Zùlpicherstr. 77,  
50937, Köln, Germany

<sup>2</sup> Joseph Henry Laboratories of Physics and Lewis-Sigler Institute for Integrative Genomics,  
Princeton University, Princeton, NJ 08544, USA

## Abstract

Molecular phenotypes link genomic information with organismic functions, fitness, and evolution. Quantitative traits are complex phenotypes that depend on multiple genomic loci. In this paper, we study the adaptive evolution of a quantitative trait under time-dependent selection, which arises from environmental changes or through fitness interactions with other co-evolving phenotypes. We analyze a model of trait evolution under mutations and genetic drift in a single-peak fitness seascape. The fitness peak performs a constrained random walk in the trait amplitude, which determines the time-dependent trait optimum in a given population. We derive analytical expressions for the distribution of the time-dependent trait divergence between populations and of the trait diversity within populations. Based on this solution, we develop a method to infer adaptive evolution of quantitative traits. Specifically, we show that the ratio of the average trait divergence and the diversity is a universal function of evolutionary time, which predicts the stabilizing strength and the driving rate of the fitness seascape. From an information-theoretic point of view, this function measures the macro-evolutionary entropy in a population ensemble, which determines the predictability of the evolutionary process. Our solution also quantifies two key characteristics of adapting populations: the cumulative fitness flux, which measures the total amount of adaptation, and the adaptive load, which is the fitness cost due to a population's lag behind the fitness peak.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Evolutionary dynamics of quantitative traits</b>	<b>3</b>
2.1	Diffusion equations for trait mean and diversity . . . . .	3
2.2	Stochastic seascape models . . . . .	6
2.3	Joint dynamics of trait and selection . . . . .	8
<b>3</b>	<b>Adaptive evolution in a single-peak fitness seascape</b>	<b>9</b>
3.1	Stationary distribution of mean and optimal trait . . . . .	9
3.2	Time-dependent trait divergence . . . . .	11
3.3	Stationary trait diversity . . . . .	14
<b>4</b>	<b>Fitness and entropy of adaptive processes</b>	<b>16</b>
4.1	Genetic load . . . . .	17
4.2	Fitness flux . . . . .	17
4.3	Predictability and entropy production . . . . .	20
<b>5</b>	<b>Inference of adaptive trait evolution</b>	<b>20</b>
<b>6</b>	<b>Discussion</b>	<b>22</b>
<b>A</b>	<b>Analytical theory of the adaptive ensemble</b>	<b>26</b>
<b>B</b>	<b>Numerical simulations</b>	<b>28</b>

## 1 Introduction

This is the second in a series of papers on the evolution of quantitative traits in biological systems [1]. We focus on molecular traits such as protein binding affinities or gene expression levels, which are mesoscopic phenotypes that bridge between genomic information and higher-level organismic traits. Such phenotypes are complex: they depend on tens to hundreds of constitutive genomic sites and are generically polymorphic in a population. Moreover, their evolution is often a strongly correlated process that involves linkage disequilibrium, i.e., allele associations due to incomplete recombination, and epistasis, i.e., fitness interactions, between constitutive sites. Hence, the evolutionary statistics of molecular quantitative traits have to go beyond traditional quantitative genetics [2–11]. Our aim is to derive *universal* phenotypic features of these processes, which decouple from details of a trait’s genomic encoding and of the molecular evolutionary dynamics.

In this paper, we focus on the adaptive evolution of molecular traits, which involves mutations, genetic drift, and (partial) recombination of the trait loci. The adaptive dynamics take place on macro-evolutionary time scales and can generate large trait changes — in contrast to micro-evolutionary processes based on standing trait variation in a population. Adaptive trait changes are driven by time-dependent selection on the trait values. Specifically, we consider the trait evolution in a single-peak *fitness seascape* [12–14], which has a moving peak described by a stochastic process in the trait coordinate. The time-dependence of the optimal trait value can have extrinsic or intrinsic causes; for example, the optimal expression level of a gene is affected by changes in the environment of an organism and by expression changes of other genes in the same gene network. These fitness seascape models have two fundamental parameters: the *stabilizing strength* and the *driving rate*, which measure the width and the mean square displacement of the fitness peak per unit of evolutionary time.

Here, we mainly focus on *macro-evolutionary* fitness seascapes, which have low driving rates compared to the diffusion of the trait by genetic drift [12,15]. This kind of selection generates two complementary evolutionary forces. On short time scales, a single fitness peak acts as stabilizing selection, which constrains the trait diversity within a population as well as its divergence between populations. On longer time scales, the population trait mean follows the moving fitness peak, which generates an adaptive component of the trait divergence. In an ensemble of populations with independent fitness peak displacements, these dynamics describe lineage-specific adaptive pressure. We discuss specific seascape models with continuous or punctuated adaptive pressure; that is, the fitness peak performs a constrained (Ornstein-Uhlenbeck) random walk or a Poisson jump process in the trait coordinate<sup>1</sup>. These stochastic processes define minimal non-equilibrium models for the adaptive evolution of a quantitative trait. In the limit case of a static fitness landscape, we recover the evolutionary equilibrium of quantitative traits under stabilizing selection, which has been the subject of a previous publication [1]. We also discuss the opposite regime of *micro-evolutionary* fitness seascapes, in which the evolution of the trait partially decouples from the movement of the fitness peak.

Our model of adaptive trait evolution contains different sources of stochasticity: mutations establish trait differences between individuals within one population, reproductive fluctuations (genetic drift) and fitness seascape fluctuations generate trait differences between populations with time. In macro-evolutionary fitness seascapes, these stochastic forces act on different time scales and define different statistical ensembles, similar to thermal and quenched fluctuations in the statistical thermodynamics of disordered systems. In section 2, we derive stochastic evolution equations for the trait mean, the trait diversity, and the position of the fitness peak, which establish a joint dynamical model for the trait and the underlying fitness seascape over macro-evolutionary time-scales. In section 3, we discuss the analytical solution of these models for a stationary ensemble of adapting populations. This ensemble has a time-independent trait diversity within populations, as well as a trait divergence between populations that depends on their divergence time. In section 4, we evaluate two important summary statistics of adaptive processes. The *genetic load*, which is defined as the difference between the maximum fitness and the mean population fitness, is shown to include a specific adaptive component, which results from the lag of the population behind the moving fitness peak. The *cumulative fitness flux* measures the amount of adaptation in a population over a macro-evolutionary period: it is zero at evolutionary equilibrium and increases monotonically with the driving rate of selection [17]. Furthermore, we determine the *predictability* of trait values in one population given its distribution in another population, which is given by a suitably defined entropy of the population ensemble under divergent evolution.

The statistical theory of this paper provides a new method to infer selection on a quantitative trait from diversity and time-resolved divergence data. Given these data in a family of evolving populations, we use the divergence-diversity ratio  $\Omega(\tau)$  for different divergence times  $\tau$  to determine stabilizing strength and driving rate of the underlying fitness seascape. These selection parameters, in turn, quantify the amount of conservation and adaptation in the evolution of the trait. Unlike previously used measures of trait evolution [16,18], the divergence-diversity ratio is universal: it is determined by stabilizing strength and driving rate of the fitness seascape and on the evolutionary distance between populations, but it depends only weakly on the trait’s constitutive sites, on the amount of recombination between these sites, and on details of the fitness dynamics. In contrast to most sequence evolution tests, the  $\Omega$  test does not require the gauge of a neutrally evolving “null trait”. We discuss this test statistics in section 5. For the reader not interested in technical details, Figure 1 provides a fast track through the preceding sections.

The statistical theory of this paper provides a new method to infer selection on a quantitative trait from diversity and time-resolved divergence data. Given these data in a family of evolving populations, we use the divergence-diversity ratio  $\Omega(\tau)$  for different divergence times  $\tau$  to determine the stabilizing strength and the driving rate of the underlying fitness seascape. These selection parameters, in turn, quantify the amount of conservation and adaptation in the evolution of the trait. Unlike previously used measures of trait evolution [16,18], the divergence-diversity ratio is universal: it depends on the stabilizing strength and the driving rate of the fitness seascape as well as on the evolutionary distance between populations, but it

---

<sup>1</sup>The stochastic process of the fitness peak should not be confused with an Ornstein-Uhlenbeck dynamics of the trait mean that describes evolutionary equilibrium in a quadratic fitness landscape [16].

is largely independent of the trait’s constitutive sites, of the amount of recombination between these sites, and of the details of the fitness dynamics. In contrast to most sequence evolution tests, the  $\Omega$  test does not require the gauge of a neutrally evolving “null trait”. We discuss this test statistics in section 5. For the reader not interested in technical details, Figure 1 provides a fast track through the preceding sections.

## 2 Evolutionary dynamics of quantitative traits

In this section, we develop minimal models for the adaptive evolution of quantitative traits in a fitness seascape. We first review the diffusion dynamics for the trait mean and the diversity under genetic drift and mutations in a given fitness landscape, which have been derived in a previous paper [1]. Second, we introduce simple stochastic models for the dynamics of selection, which promote fitness landscapes to fitness seascapes. We then combine the dynamics of trait and selection to a joint, non-equilibrium evolutionary model.

### 2.1 Diffusion equations for trait mean and diversity

Our model for quantitative traits is based on a simple additive map from genotypes to phenotypes. The trait value  $E$  of an individual depends on its genotype  $\mathbf{a} = (a_1, \dots, a_\ell)$  at  $\ell$  constitutive genomic sites,

$$E(\mathbf{a}) = \sum_{i=1}^{\ell} E_i \sigma_i, \quad \text{with } \sigma_i \equiv \begin{cases} 1, & \text{if } a_i = a_i^*, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Here, the trait is measured from its minimum value, and  $E_i > 0$  is the contribution of a given site  $i$  to the trait value. We assume a two-allele genomic alphabet and  $a_i^*$  denotes the allele conferring the larger phenotype at site  $i$ . The extension to a four-allele alphabet is straightforward. The genotype-phenotype map (1) defines the allelic trait average  $\Gamma_0$  and the trait span  $E_0^2$ ,

$$\Gamma_0 = \frac{1}{2} \sum_{i=1}^{\ell} E_i, \quad E_0^2 = \frac{1}{4} \sum_{i=1}^{\ell} E_i^2. \quad (2)$$

Quantitative traits have a sufficient number of constitutive loci to be generically polymorphic in a population, although most individual genomic sites are monomorphic. The distribution of trait values in a given population,  $\mathcal{W}(E)$ , is often approximately Gaussian [1, 3, 14]. Hence, it is well characterized by its mean and variance,

$$\begin{aligned} \Gamma \equiv \overline{E} &= \int dE E \mathcal{W}(E), \\ \Delta \equiv \overline{(E - \Gamma)^2} &= \int dE (E - \Gamma)^2 \mathcal{W}(E), \end{aligned} \quad (3)$$

where overbars denote averages over the trait distribution  $\mathcal{W}(E)$  within a population. The variance  $\Delta$  will be called the trait diversity; in the language of quantitative genetics, this quantity equals the total heritable variance including epistatic effects.

We consider the evolution of the trait  $E$  under genetic drift, genomic mutations, and natural selection, which is given by a trait-dependent fitness seascape  $f(E, t)$  that changes on macro-evolutionary time scales. This process is illustrated in Fig. 1: At a given evolutionary time, the trait distribution in a population has mean  $\Gamma(t)$ , diversity  $\Delta(t)$ , and is positioned at a distance  $\Lambda(t) \equiv \Gamma(t) - E^*(t)$  from the optimal trait value. The trait distribution follows the moving fitness peak, building up a trait divergence

$$D^{(1)}(t, \tau) = (\Gamma(t) - \Gamma(t - \tau))^2 \quad (4)$$

between an ancestral population at time  $t - \tau$  and its descendent population at time  $t$  in a given lineage. In the same way, we can define the trait divergence between two descendent populations at time  $t$  that have evolved independently from a common ancestor population at time  $t - \tau/2$ ,

$$D^{(2)}(t, \tau) = (\Gamma_1(t) - \Gamma_2(t))^2. \quad (5)$$

In a suitably defined ensemble of parallel-evolving populations, the expectation values of these divergences,  $\langle D^{(\kappa)}(\tau) \rangle$  ( $\kappa = 1, 2$ ), depend only on the divergence time  $\tau$ . The asymptotic divergence for long times is just twice the trait variance across populations,

$$\lim_{\tau \rightarrow \infty} \langle D^{(\kappa)}(\tau) \rangle = 2 \langle (\Gamma - \langle \Gamma \rangle)^2 \rangle \quad (\kappa = 1, 2). \quad (6)$$

In particular, the quantity  $E_0^2$  defined in (2) is the trait variance in an ensemble of random genotypes, which results from neutral evolution (with averages  $\langle \dots \rangle_0$  marked by a subscript) at low mutation rates,  $E_0^2 = \lim_{\mu \rightarrow 0} \langle (\Gamma - \langle \Gamma \rangle_0)^2 \rangle_0$ . For finite times, however, the statistics of the single-lineage divergence  $D^{(1)}$  and the cross-lineage divergence  $D^{(2)}$  differ from each other in an adaptive process. As we will discuss in detail below, this is a manifestation of the non-equilibrium evolutionary dynamics in a fitness seascape. In contrast, evolutionary equilibrium in a fitness landscape dictates  $\langle D^{(1)}(\tau) \rangle_{\text{eq}} = \langle D^{(2)}(\tau) \rangle_{\text{eq}}$  by detailed balance.

As shown in a previous companion paper [1], the evolutionary dynamics of a quantitative trait in a fitness seascape can be described in good approximation by diffusion equations for its mean and its diversity,

$$\frac{\partial}{\partial t} Q(\Gamma, t | F_1) = \left[ \frac{g^{\Gamma\Gamma}}{2N} \frac{\partial^2}{\partial \Gamma^2} - \frac{\partial}{\partial \Gamma} \left( m^\Gamma + g^{\Gamma\Gamma} \frac{\partial F_1(\Gamma, t)}{\partial \Gamma} \right) \right] Q(\Gamma, t | F_1), \quad (7)$$

$$\frac{\partial}{\partial t} Q(\Delta, t | F_2) = \left[ \frac{g^{\Delta\Delta}}{2N} \frac{\partial^2}{\partial \Delta^2} - \frac{\partial}{\partial \Delta} \left( m^\Delta + g^{\Delta\Delta} \frac{\partial F_2(\Delta, t)}{\partial \Delta} \right) \right] Q(\Delta, t | F_2), \quad (8)$$

which are projections of the Kimura diffusion equation [19,20] from genotypes onto the phenotype space. The distributions  $Q(\Gamma, t | F_1)$  and  $Q(\Delta, t | F_2)$  describe an ensemble of populations evolving in the same fitness seascape  $f(E, t)$ . These dynamics involve the fitness seascape components

$$F_1(\Gamma, t) = f(\Gamma, t) + f''(\Gamma, t) \times \int d\Delta \Delta Q_2(\Delta, t | F_2), \quad (9)$$

$$F_2(\Delta, t) = \Delta \times \int d\Gamma f''(\Gamma, t) Q(\Gamma, t | F_1), \quad (10)$$

which are projections of the mean population fitness

$$\bar{f}(t) \equiv \int dE f(E, t) \mathcal{W}(E, t) = f(\Gamma, t) + \frac{1}{2} \Delta f''(\Gamma, t) + \dots \quad (11)$$

onto the marginal variables  $\Gamma$  and  $\Delta$ . Genetic drift enters through the diffusion coefficients  $g^{\Gamma\Gamma} = \langle \Delta \rangle \equiv \int d\Delta \Delta Q(\Delta, t | F_2)$  and  $g^{\Delta\Delta} = 2\Delta^2$ , mutations through the mutation coefficients  $m^\Gamma = -2\mu(\Gamma - \Gamma_0)$  and  $m^\Delta = -4\mu(\Delta - E_0^2) - \Delta/N$ ; these coefficients depend on the effective population size  $N$  and the point mutation rate  $\mu$ . The diffusion equations (7) and (8) are coupled through the fitness components (9) and (10) and through the diffusion coefficient  $g^{\Gamma\Gamma}$ . In the absence of direct selection on the trait mean, i.e. for  $F_1(\Gamma, t) = 0$ , eq. (7) describes a “quasi-neutral” diffusion of the trait mean, which depends the full drift term  $g^{\Gamma\Gamma} = \langle \Delta \rangle(c)$  under selection. This dynamics defines a characteristic time scale

$$\tilde{\tau} \equiv \frac{2NE_0^2}{\langle \Delta \rangle}. \quad (12)$$

Recombination between the trait loci induces a crossover between selection on entire genotypes and selection on individual alleles [21,22]. This affects the form of the diffusive dynamics of  $\Delta$  and, hence, the coefficients

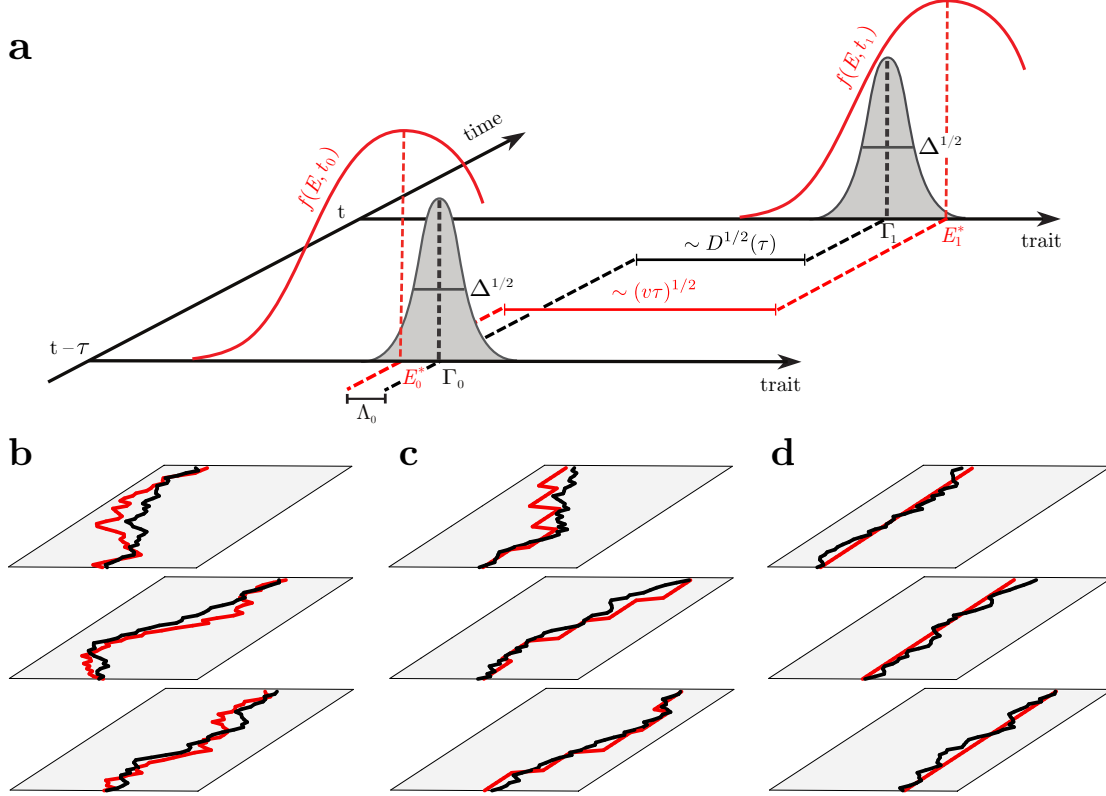


Figure 1: **Adaptive evolution of a quantitative trait.** (a) Evolution of the distribution of trait values  $\mathcal{W}(E, t)$  (gray curves) in a given population subject to a single-peak fitness seascape  $f(E, t)$  (red curves). At a given time  $t$ , the population has a trait distribution  $\mathcal{W}(E, t)$  with mean  $\Gamma(t)$  and diversity  $\Delta(t)$ , and is positioned at a distance  $\Lambda(t) = \Gamma(t) - E^*(t)$  from the fitness peak. The population follows the moving fitness peak and builds up a trait divergence  $D^{(1)}(t, \tau) = (\Gamma(t) - \Gamma(t - \tau))^2$  between the ancestral state at time  $t - \tau$  and the descendent state at time  $t$ . The divergence  $D^{(2)}(t, \tau)$  between two descendent populations with a common ancestor at time  $t - \tau/2$  can be defined in an analogous way; see eqs. (4) and (5). (b–d) Evolutionary population ensembles, each represented by three sample populations. In a given population, a realization of a single-peak fitness seascape specifies a lineage-specific optimal trait value that depends on evolutionary time,  $E^*(t)$  (red line). The population mean trait,  $\Gamma(t)$  (black line), adapts to the moving fitness peak with additional lineage-specific fluctuations due to mutations and genetic drift. The adaptive process is shown for three cases of fitness seascapes: (b) Diffusive fitness seascape: Incremental changes in the optimal trait value reflect adaptive pressure caused by continuous ecological changes. The function  $E^*(t)$  follows an Ornstein-Uhlenbeck random walk in the trait coordinate. (c) Punctuated fitness seascape: Sudden changes in the optimal trait value reflect adaptive pressure caused by major, discrete ecological events. The function  $E^*(t)$  is described by a Poisson jump process in the trait coordinate. We show that both types of fitness seascapes lead to a solvable, non-equilibrium joint statistics of  $\Gamma$  and  $E^*$ . (d) Fitness landscape: Each population has a time-independent optimal trait value  $E^*$  and reaches an evolutionary (selection-mutation-drift) equilibrium [1].

of the diffusion equation for  $\Gamma$ . However, these changes have only a small effect on the statistics of  $\Gamma$  over a wide range of evolutionary parameters [1].

In the special case of a time-independent fitness landscape  $f(E)$ , the diffusive dynamics of trait mean and diversity leads to evolutionary equilibria of a Boltzmann form [1],

$$Q_{\text{eq}}(\Gamma | F_1) = \frac{1}{Z_\Gamma} \tilde{Q}_0(\Gamma) \exp[2NF_1(\Gamma)], \quad (13)$$

$$Q_{\text{eq}}(\Delta | F_2) = \frac{1}{Z_\Delta} Q_0(\Delta) \exp[2NF_2(\Delta)], \quad (14)$$

where  $Z_\Gamma$  and  $Z_\Delta$  are normalization constants. The equilibrium distributions under selection build on the quasi-neutral distribution of the trait mean,  $\tilde{Q}_0(\Gamma) \sim \exp[-2\mu N(\Gamma - \Gamma_0)^2/\langle\Delta\rangle]$ , and on the neutral diversity distribution  $Q_0(\Delta)$  (see also section 3.3). We note that evolutionary equilibrium in a static fitness landscape is limited to the marginal distributions  $Q_{\text{eq}}(\Gamma | F_1)$  and  $Q_{\text{eq}}(\Delta | F_2)$ , while the joint distribution  $Q(\Gamma, \Delta | f)$  reaches a non-equilibrium stationary state [1]. In the limit of low mutation rates, the Boltzmann distribution (13) describes an asymptotic selection-drift equilibrium  $Q_{\text{eq}}(E | F_1) \sim Q_0(E) \exp[2NF(E)]$ ; the trait values  $E$  are predominantly monomorphic in a population and they change by substitutions at individual trait loci [1, 23, 24].

## 2.2 Stochastic seascape models

For a generic fitness seascape  $f(E, t)$ , the diffusion equations (7) and (8) do not have a closed analytical solution. At the same time, we are often not interested in the detailed history of fitness peak displacements and the resulting trait changes. To describe generic features of adaptive processes, we now introduce solvable stochastic models of the seascape dynamics and link broad features of these models to statistical observables of adapting populations.

In this paper, we restrict our analysis to single-peak fitness seascapes of the form

$$f(E, t) = f^* - c_0(E - E^*(t))^2. \quad (15)$$

Despite its simple form, the fitness function (15) covers a broad spectrum of interesting selection scenarios [25]. For constant trait optimum  $E^*$ , it is a time-honored model of *stabilizing selection*. [2, 6, 10, 14, 23, 25–29]. Nearly all known examples of empirical fitness landscapes for molecular quantitative traits are of single-peak [30] or mesa-shaped [26, 31–34] forms. Mesa landscapes describe directional selection with diminishing return: they contain a fitness flank on one side of a characteristic “rim” value  $E^*$  and flatten to a plateau of maximal fitness on the other side. Furthermore, trait values on the fitness plateau tend to be encoded by far fewer genotypes than low-fitness values. This differential coverage of the genotype-phenotype map turns out to generate an effective second flank of the fitness landscape, which makes our subsequent theory applicable to mesa landscapes as well [25]. We refer to the scaled parameter

$$c = 2NE_0^2 c_0 \quad (16)$$

as the *stabilizing strength* of a fitness landscape. This dimensionless quantity has a simple interpretation: it equals the ratio of the neutral trait variance  $E_0^2$  and the weakly deleterious trait variance around the fitness peak, which, by definition, produces a fitness drop  $\leq 1/(2N)$  below the maximum  $f^*$ . As shown in ref. [1], the mutation-selection-drift dynamics of a quantitative trait in a single-peak fitness landscape leads to evolutionary equilibrium with a characteristic equilibration time

$$\tau_{\text{eq}}(c) = \frac{1}{\mu + c\tilde{\tau}^{-1}(c)} \simeq \begin{cases} \mu^{-1} & \text{for } c \lesssim 1, \\ (c\tilde{\tau}(c))^{-1} & \text{for } c \gtrsim 1, \end{cases} \quad (17)$$

where  $\tilde{\tau}(c)$  is the quasi-neutral drift time defined in eq. (12).

For time-dependent  $E^*(t)$ , eq. (15) becomes a fitness seascape model [12–14]. At any given evolutionary time, this model describes stabilizing selection of strength  $c$  towards an optimal trait value  $E^*(t)$ . In

addition, the changes of  $E^*(t)$  over macro-evolutionary periods introduce directional selection on the trait and generate adaptive evolution. The form (15) of a fitness seascape assumes the stabilizing strength  $c$  to remain constant over time. As discussed in section 2.3, this assumption leads to an important computational simplification: only the trait mean  $\Gamma$  adapts to the moving fitness peak, while the diversity  $\Delta$  remains at evolutionary equilibrium. However, generalizing of our model to a time-dependent stabilizing strength  $c(t)$  is straightforward and is briefly discussed below. We consider two minimal models of seascape dynamics:

**Diffusive fitness seascapes.** In this model, the fitness optimum  $E^*(t)$  performs an Ornstein-Uhlenbeck random walk with diffusion constant  $v_0$ , average value  $\mathcal{E}$  and stationary mean square deviation  $r_0^2$ . The scaled parameters

$$v = \frac{v_0}{E_0^2}, \quad r^2 = \frac{r_0^2}{E_0^2}, \quad (18)$$

will be called the *driving rate* and the *driving span* of a fitness seascape. Different realizations of this random walk with the same set of parameters are shown in Fig. 1(b). The distribution of optimum trait values,  $R(E^*, t)$ , follows a diffusion equation,

$$\frac{\partial}{\partial t} R(E^*, t) = v E_0^2 \frac{\partial}{\partial E^*} \left[ \frac{\partial}{\partial E^*} + \frac{1}{r^2 E_0^2} (E^* - \mathcal{E}) \right] R(E^*, t). \quad (19)$$

This dynamics leads to a seascape ensemble, which is characterized by an expected peak divergence

$$\langle (E^*(t) - E^*(t + \tau))^2 \rangle = 2r^2 E_0^2 \left( 1 - e^{-\tau/\tau_{\text{sat}}(v, r^2)} \right) \quad (20)$$

with the saturation time

$$\tau_{\text{sat}}(v, r^2) = \frac{r^2}{v}, \quad (21)$$

and by an equilibrium distribution

$$R_{\text{eq}}(E^*) = \frac{1}{\sqrt{2\pi r^2 E_0^2}} \exp \left[ -\frac{1}{2} \frac{(E^* - \mathcal{E})^2}{r^2 E_0^2} \right] \quad (22)$$

of optimal trait values. Diffusive seascapes models of this form describe continuous adaptive pressure due to incremental ecological changes that affect the optimal trait value  $E^*(t)$ . We assume that typical optimal trait values fall into the neutral trait repertoire given by eq. (2), which implies that the scaled driving span  $r^2$  is at most of order 1.

**Punctuated fitness seascapes.** In this model, the fitness optimum performs a Poisson jump process with jump rate  $\tau_{\text{sat}}^{-1}(v, r^2) = v/r^2$ , by which successive values of  $E^*$  are drawn independently from the distribution  $R_{\text{eq}}(E^*)$ , given by (22). Different realizations of this process are shown in Fig. 1(c). Fitness jumps may result from discrete ecological events such as major migrations or speciations. The Poisson jump process is described by the evolution equation

$$\frac{\partial}{\partial t} R(E^*, t) = \frac{v}{r^2} [R_{\text{eq}}(E^*) - R(E^*, t)]. \quad (23)$$

It has the same time-dependent expected peak divergence (20) and the same equilibrium distribution (22) as the diffusion process (19) with same driving parameters (18). The difference between the jump process and the diffusion process lies in the anomalous scaling of higher moments,

$$\langle (E^*(t) - E^*(t + \tau))^k \rangle \sim E_0^k r^{k-2} v \tau \quad \text{for } k = 2, 4, \dots \text{ and } \tau \ll \tau_{\text{sat}}(v, r^2). \quad (24)$$

This scaling is shared by simple models of turbulence; see, e.g., ref. [35].

In both types of fitness seascape, we distinguish two dynamical selection regimes:



- *Macro-evolutionary* fitness seascapes are defined by the condition  $\tau_{\text{sat}}(v, r^2) \gtrsim \tau_{\text{eq}}(c)$ . As discussed in detail below, such seascapes keep the trait mean always close to equilibrium and induce an adaptive response linear in the driving rate  $v$ . The limit  $v \rightarrow 0$  describes an ensemble of *quenched* population-specific fitness *landscapes* with a distribution of optimal trait values given by eq. (22); see Fig. 1(d).
- *Micro-evolutionary* fitness seascapes have  $\tau_{\text{sat}}(v, r^2) \lesssim \tau_{\text{eq}}(c)$  and delineate a regime of reduced adaptive response, where the evolution of the trait mean gradually decouples from that of the fitness seascape. In the asymptotic fast-driving regime ( $v \gg r^2/\tau_{\text{eq}}(c)$ ), the adaptation of the trait is completely suppressed. In this regime, we can average over the fitness fluctuations and describe the macro-evolution of the trait in terms of an effective fitness landscape with an optimal trait value  $\mathcal{E}$ .

### 2.3 Joint dynamics of trait and selection

We now combine the diffusive dynamics of quantitative traits in a given fitness seascape, which is given by eqs. (7) and (8), and the seascape dynamics (19) or (23) into a stochastic model of adaptive evolution. The statistical ensemble generated by this model is illustrated in Fig. 1(b-d): Each population evolves in a specific realization of the fitness seascape, which is given by a history of peak values  $E^*(t)$ . Its trait mean  $\Gamma(t)$  follows the moving fitness peak with fluctuations due to mutations and genetic drift. The ensemble of populations contains, in addition, the stochastic differences between realizations of the fitness seascape. The statistics of this ensemble involves combined averages over both kinds of fluctuations, which are denoted by angular brackets  $\langle \dots \rangle$ .

The population ensemble can be described by a joint distribution of mean and optimum trait values,  $Q(\Gamma, E^*, t) = Q(\Gamma, t | E^*)R(E^*, t)$ . Using eqs. (7) and (19) together with the projection of the fitness seascape,

$$F_1(\Gamma | E^*) = f^* - \frac{c}{NE_0^2} \langle \Delta \rangle - \frac{c}{E_0^2} (\Gamma - E^*)^2, \quad (25)$$

given by eqs. (9) and (15), we obtain the evolution equation for the joint distribution in a diffusive seascape,

$$\frac{\partial}{\partial t} Q(\Gamma, E^*, t) = \left[ \frac{g^{\Gamma\Gamma}}{2N} \frac{\partial^2}{\partial \Gamma^2} - \frac{\partial}{\partial \Gamma} \left( m^\Gamma - g^{\Gamma\Gamma} \frac{2c}{E_0^2} (\Gamma - E^*) \right) + \frac{v}{E_0^2} \frac{\partial^2}{\partial E^{*2}} + \frac{v}{r^2} \frac{\partial}{\partial E^*} (E^* - \mathcal{E}) \right] Q(\Gamma, E^*, t), \quad (26)$$

with  $g^{\Gamma\Gamma} = \langle \Delta \rangle$  and  $m^\Gamma = -2\mu(\Gamma - \Gamma_0)$ . Note that the differential operator in eq. (26) is asymmetric: the trait optimum  $E^*$  follows an independent stochastic dynamics, but the trait mean  $\Gamma$  is coupled to  $E^*$ . This asymmetry reflects the causal relation between selection and adaptive response: the trait mean  $\Gamma(t)$  follows the moving fitness peak, as shown in Fig. 1(b,c). As a consequence, the joint evolution equation (26) leads to a non-equilibrium stationary distribution  $Q_{\text{stat}}(\Gamma, E^*)$ , although the marginal seascape dynamics (19) reaches an equilibrium state. Only in the fitness landscape limit ( $v \rightarrow 0$ ), the evolution of the trait mean reaches evolutionary equilibrium. In the next section, we will obtain explicit solutions for the non-equilibrium distribution  $Q_{\text{stat}}(\Gamma, E^*)$  and its equilibrium limit. The case of a punctuated fitness seascape is treated in Appendix A, where we solve the Langevin equations for  $\Gamma$  and  $E^*$  to obtain the first and second moments of  $Q(\Gamma, E^*, t)$ .

The trait diversity evolves under the projected fitness function

$$F_2(\Delta | c) = -\frac{c}{NE_0^2} \Delta, \quad (27)$$

given by eqs. (15) and (10). In a fitness seascapes with a constant stabilizing strength  $c$ , this function is time-independent. The dynamics of the trait diversity (8) decouples from the adaptive evolution of the trait mean and leads an evolutionary equilibrium  $Q_{\text{eq}}(\Delta | c)$  of the form (14). As detailed in Section 3.3, the equilibrium assumption for the trait diversity holds for most adaptive processes in a fitness seascape of the form (15). However, we can generalize our seascape models to include a time-dependent stabilizing strength  $c(t)$ . This leads to generic adaptive evolution of both,  $\Gamma$  and  $\Delta$ , which is described by a coupled non-equilibrium stationary distribution  $Q_{\text{stat}}(\Gamma, \Delta, E^*, c)$ .

### 3 Adaptive evolution in a single-peak fitness seascape

In this section, we develop the key analytical results of this paper. We provide an explicit solution for the non-equilibrium joint distribution of mean and optimal trait in a diffusive seascape,  $Q_{\text{stat}}(\Gamma, E^*)$ ; the case of punctuated seascapes is treated in Appendix A. These solutions describe a stationary ensembles of adapting populations. We derive an expression for the expected time-dependent trait divergence in these ensembles, which holds for both seascape models. Finally, we juxtapose the adaptive behavior of the trait mean with the equilibrium statistics of the trait diversity, which emerges in good approximation for most fitness seascape of constant stabilizing strength. Our analytical results are supported by simulations for diffusive and punctuated fitness seascapes.

#### 3.1 Stationary distribution of mean and optimal trait

In a diffusive fitness seascape, the evolution equation (26) has a stationary solution of bivariate Gaussian form,

$$Q_{\text{stat}}(\Gamma, E^*) = \frac{1}{Z} \exp \left[ -\frac{1}{2} \begin{pmatrix} \hat{\Gamma} \\ \hat{E}^* \end{pmatrix}^T \Sigma^{-1} \begin{pmatrix} \hat{\Gamma} \\ \hat{E}^* \end{pmatrix} \right], \quad (28)$$

with the expectation values

$$\begin{aligned} \begin{pmatrix} \langle \Gamma \rangle \\ \langle E^* \rangle \end{pmatrix} &\equiv \int d\Gamma dE^* \begin{pmatrix} \Gamma \\ E^* \end{pmatrix} Q_{\text{stat}}(\Gamma, E^*) \\ &= \begin{pmatrix} w(c) \mathcal{E} + (1 - w(c)) \Gamma_0 \\ \mathcal{E} \end{pmatrix}, \end{aligned} \quad (29)$$

and the covariance matrix

$$\begin{aligned} \Sigma = \begin{pmatrix} \langle \hat{\Gamma}^2 \rangle & \langle \hat{\Gamma} \hat{E}^* \rangle \\ \langle \hat{\Gamma} \hat{E}^* \rangle & \langle \hat{E}^{*2} \rangle \end{pmatrix} &\equiv \int d\Gamma dE^* \begin{pmatrix} \hat{\Gamma}^2 & \hat{\Gamma} \hat{E}^* \\ \hat{\Gamma} \hat{E}^* & \hat{E}^{*2} \end{pmatrix} Q_{\text{stat}}(\Gamma, E^*) \\ &= E_0^2 \begin{pmatrix} (1/2c) w(c) + r^2 w(c) w(c, v, r^2) & r^2 w(c, v, r^2) \\ r^2 w(c, v, r^2) & r^2 \end{pmatrix}, \end{aligned} \quad (30)$$

where  $\hat{\Gamma} \equiv \Gamma - \langle \Gamma \rangle$  and  $\hat{E}^* \equiv E^* - \mathcal{E}$ . The stationary distribution  $Q_{\text{stat}}(\Gamma, E^*)$  depends on the parameters that characterize the fitness seascape: the stabilizing strength  $c$ , the driving rate  $v$ , and the relative driving span  $r^2$ , which are defined in eqs. (16) and (18). Together with the effective population size  $N$  and the point mutation rate  $\mu$ , these parameters determine the characteristic time scales of evolution in a fitness seascape, the equilibration time  $\tau_{\text{eq}}(c)$  and the saturation time of fitness fluctuations,  $\tau_{\text{sat}}(v, r^2)$ ; see eqs. (17) and (21). The function

$$w(c, v, r^2) \equiv \frac{c \langle \delta \rangle}{c \langle \delta \rangle + 2\theta + Nv/r^2} = \frac{\tau_{\text{eq}}^{-1}(c) - \mu}{\tau_{\text{eq}}^{-1}(c) - \mu + \tau_{\text{sat}}^{-1}(v, r^2)}, \quad (31)$$

and its equilibrium limit  $w(c) \equiv w(c, v=0, r^2)$  govern the coupling between the mean and optimal trait. The mutation rate  $\mu$  is the inverse of the neutral timescale  $\tau_{\text{eq}}(0) = \mu^{-1}$ . These functions depend on the scaled diversity  $\langle \delta \rangle \equiv \langle \Delta \rangle / E_0^2$ , which is given in ref. [1], eqs. (68) – (73), and is restated below in eq. (49). For traits under substantial selection ( $c \gtrsim 1$ ), we can distinguish two dynamical regimes: In macro-evolutionary fitness seascapes, where  $\tau_{\text{sat}}(v, r^2) \gtrsim \tau_{\text{eq}}(c) \approx 2N/(\langle \delta \rangle c)$ , this coupling remains close to the equilibrium value  $w(c) \approx 1$ ; micro-evolutionary fitness fluctuations, which have  $\tau_{\text{sat}}(v, r^2) \lesssim \tau_{\text{eq}}(c)$ , induce a partial decoupling of mean and optimal trait.

We can also express this crossover in terms of the average square distance between trait mean in the population and optimal trait of the underlying fitness seascape,

$$\langle \Lambda^2 \rangle \equiv \int d\Gamma dE^* (\Gamma - E^*)^2 Q_{\text{stat}}(\Gamma, E^*). \quad (32)$$

The analytical solution for the scaled quantity  $\langle \lambda^2 \rangle \equiv \langle \Lambda^2 \rangle / E_0^2$ , follows from eqs. (29) and (30),

$$\langle \lambda^2 \rangle(c, v, r^2) \simeq \begin{cases} \langle \lambda^2 \rangle_{\text{eq}}(c, r^2) + v \tau_{\text{eq}}(c) \frac{w(c)}{2} \left[ 1 + \mathcal{O}\left(\frac{\tau_{\text{eq}}}{\tau_{\text{sat}}}\right) \right], & (\text{macroev. seascapes}) \\ \langle \lambda \rangle_{\text{eq}}(c, 0) + r^2 \left[ 1 - \mathcal{O}\left(\frac{\tau_{\text{sat}}}{\tau_{\text{eq}}}\right) \right], & (\text{microev. seascapes}), \end{cases} \quad (33)$$

where

$$\begin{aligned} \langle \lambda^2 \rangle_{\text{eq}}(c, r^2) &= \frac{w(c)}{2c} + (\langle \lambda^2 \rangle_0 + r^2)(1 - w(c))^2 \\ &\simeq \frac{1}{2c} \quad \text{for } c \gg 1 \end{aligned} \quad (34)$$

is the equilibrium average in a fitness landscape. In micro-evolutionary seascapes, this distance remains small, which indicates that the trait distribution  $\mathcal{W}(E)$  efficiently follows the displacements of the fitness peak. In macro-evolutionary seascapes, the mean square distance  $\langle \lambda^2 \rangle$  becomes comparable to the driving span  $r^2$ ; that is, the population no longer follows the moving fitness peak in an efficient way.

The distribution  $Q_{\text{stat}}(\Gamma, E^*)$  describes a stationary state that is manifestly out of equilibrium, i.e., it does not have detailed balance. Its probability current

$$\begin{aligned} \mathbf{J}_{\text{stat}}(\Gamma, E^*) &= - \left( \frac{g^{\Gamma\Gamma}}{2N} \frac{\partial}{\partial \Gamma} - m^\Gamma + g^{\Gamma\Gamma} \frac{2c}{NE_0^2} (\Gamma - E^*) \right) Q_{\text{stat}}(\Gamma, E^*) \\ &\quad - \left( v \frac{\partial}{\partial E^*} + \frac{v}{r^2} (E^* - \mathcal{E}) \right) Q_{\text{stat}}(\Gamma, E^*) \\ &\simeq \begin{cases} \left[ -2vc \left( \frac{\hat{\Gamma} - \hat{E}^*(1 + 1/(2cr^2))}{(\hat{\Gamma} - \hat{E}^*)} \right) \left( 1 + \mathcal{O}\left(\frac{\tau_{\text{eq}}}{\tau_{\text{sat}}}\right) \right) \right] Q_{\text{stat}}(\Gamma, E^*), & (\text{macroev. seascapes}) \\ \left[ \frac{c\langle \delta \rangle}{N} \left( \frac{\hat{E}^*}{-2cr^2 \hat{\Gamma}} \right) \left( 1 - \mathcal{O}\left(\frac{\tau_{\text{sat}}}{\tau_{\text{eq}}}\right) \right) \right] Q_{\text{stat}}(\Gamma, E^*) & (\text{microev. seascapes}), \end{cases} \end{aligned} \quad (35)$$

expresses the adaptive motion of the trait mean following the displacements of the fitness peak. The probability current shows a crossover similar to the adaptive part of  $\langle \Lambda^2 \rangle$  in (33): it increases linearly for low driving rates and saturates to a constant in the regime of micro-evolutionary fitness fluctuations.

Remarkably, the joint statistics of mean and optimal trait can be associated with evolutionary equilibrium in the limits of low and high driving rates. In the first case, we obtain the equilibrium distribution

$$\begin{aligned} Q_{\text{eq}}(\Gamma, E^*) &= \lim_{v \rightarrow 0} Q_{\text{stat}}(\Gamma, E^*) \\ &= \tilde{Q}_0(\Gamma) \exp[2NF_1(\Gamma|E^*)] R(E^*) \\ &= \frac{1}{Z_\Gamma \sqrt{2\pi r^2 E_0^2}} \exp \left[ -\frac{2\theta}{\langle \Delta \rangle} (\Gamma - \Gamma_0)^2 - \frac{c}{E_0^2} (\Gamma - E^*)^2 - \frac{1}{2r^2 E_0^2} (E^* - \mathcal{E})^2 \right], \end{aligned} \quad (36)$$

which is the product of a Boltzmann distribution (13) and a *quenched* weight of the trait optimum  $E^*$  given by (22). This distribution satisfies detailed balance; that is, the probability current  $\mathbf{J}_{\text{stat}}(\Gamma, E^*)$  vanishes in the limit  $v \rightarrow 0$ . In the opposite limit, we obtain the distribution

$$\begin{aligned} Q_\infty(\Gamma, E^*) &= \lim_{v \rightarrow \infty} Q_{\text{stat}}(\Gamma, E^*) \\ &= \tilde{Q}_0(\Gamma) \exp[2NF_1(\Gamma|\mathcal{E})] R(E^*) \\ &= \frac{1}{Z_\Gamma \sqrt{2\pi r^2 E_0^2}} \exp \left[ -\frac{2\theta}{\langle \Delta \rangle} (\Gamma - \Gamma_0)^2 - \frac{c}{E_0^2} (\Gamma - \mathcal{E})^2 - \frac{1}{2r^2 E_0^2} (E^* - \mathcal{E})^2 \right]. \end{aligned} \quad (37)$$

In this limit, the fast fluctuations of the fitness peak — and the associated current  $\mathbf{J}_{\text{stat}}(\Gamma, E^*)$  given by (35) — decouple from the macro-evolutionary dynamics of the mean trait. The latter is governed by the effective fitness landscape

$$F_1(\Gamma|\mathcal{E}) = \int F_1(\Gamma|E^*) R(E^*) dE^*, \quad (38)$$

which is obtained by averaging over the ensemble (22) of fitness peak positions and it describes stabilizing selection towards the average peak position  $\mathcal{E}$ . Accordingly, the scaled average square distance  $\langle \lambda^2 \rangle$ , as given by eq. (33), is the sum of the equilibrium variance  $\langle \lambda^2 \rangle_{\text{eq}}(c, 0)$  and the driving span  $r^2$ . We can extend the notion of an effective fitness landscape to micro-evolutionary seascapes with a large but finite driving rate ( $c \gg 1$ ,  $v \gg r^2/\tau_{\text{eq}}(c)$ ). Such seascape models still generate stabilizing selection on the trait mean towards the mean peak position  $\mathcal{E}$ , but with a reduced effective stabilizing strength

$$c_{\text{eff}} \sim c \left[ 1 - \frac{2c^2 r^2 \tau_{\text{sat}}(v, r^2)}{\tau_{\text{eq}}(c)} \right]. \quad (39)$$

Similar effective landscapes resulting from micro-evolutionary seascapes have been observed in phenomenological models [36].

As shown in Appendix A, the dynamics of the trait in a punctuated seascape leads to a stationary population ensemble that has the same first and second moments as in the case of a diffusive seascape. In particular, the average square displacement between mean and optimal trait, eq. (33), as well as the averages of divergence, genetic load, and fitness flux described in the following sections coincide for both kinds of seascapes.

The properties of the stationary ensemble of mean and optimal trait in a fitness seascape are summarized in Fig. 2. The stationary distribution  $Q_{\text{stat}}(\Gamma, E^*)$  is shown in Fig. 2(a–c) for given parameters  $c$ ,  $r^2$ , and for different values of the driving rate: in the equilibrium limit ( $v \rightarrow 0$ ), for an intermediate value of  $v$ , and in the fast-driving regime ( $v \gg r^2/\tau_{\text{eq}}(c)$ ). The non-equilibrium probability current  $\mathbf{J}_{\text{stat}}(\Gamma, E^*)$  is marked by arrows. The crossover between micro- and macro-evolutionary fitness seascapes is plotted in Fig. 2(d,e) for the scaled average square distance  $\langle \lambda^2 \rangle$  as a function of the driving rate  $v$ . Our analytical results are tested by numerical simulations of the underlying Fisher-Wright process [37] in a fitness seascape (15) with diffusive and punctuated peak displacement. The details of the numerical methods for the population simulations are discussed in Appendix B.

### 3.2 Time-dependent trait divergence

In the previous paper [1], we have shown that the variance of the trait mean across populations,  $\langle \hat{\Gamma}^2 \rangle$ , and the average trait diversity  $\langle \Delta \rangle$  uniquely characterize the stabilizing strength  $c$  in a fitness landscape. The ensemble variance  $\langle \hat{\Gamma}^2 \rangle$  is just the half of the asymptotic trait divergence  $\lim_{\tau \rightarrow \infty} \langle D(\tau) \rangle \equiv \lim_{\tau \rightarrow \infty} \langle (\Gamma(t + \tau) - \Gamma(t))^2 \rangle$ . As it is clear from the previous subsection, the stationary distribution  $Q_{\text{stat}}(\Gamma)$  and its statistics is compatible with different values of the seascape parameters and, hence, cannot uniquely characterize them. Instead, we will use the time-dependent average trait divergence to characterize the parameters of the fitness seascape. This quantity can be estimated between populations in one and two lineages,  $\langle D^{(\kappa)} \rangle(\tau)$  ( $\kappa = 1, 2$ ), as defined in eqs. (4) and (5). The average divergence between an ancestral and a descendent population in a single lineage can be written as an expectation value in the stationary ensemble,

$$\begin{aligned} \langle D^{(1)} \rangle(\tau) &\equiv \langle (\Gamma(t) - \Gamma(t_a))^2 \rangle \\ &\equiv \int d\Gamma d\Gamma_a (\Gamma - \Gamma_a)^2 \times \left( \begin{array}{c} \text{Diagram: A line segment from } \Gamma_a \text{ to } \Gamma \text{ with a point at } \Gamma \text{ and a point at } \Gamma_a. \text{ An arrow labeled } \tau \text{ points from } \Gamma_a \text{ to } \Gamma. \end{array} \right) \\ &= \int dE_a^* dE^* d\Gamma_a d\Gamma (\Gamma - \Gamma_a)^2 G_\tau(\Gamma, E^* | \Gamma_a, E_a^*) Q_{\text{stat}}(\Gamma_a, E_a^*), \end{aligned} \quad (40)$$

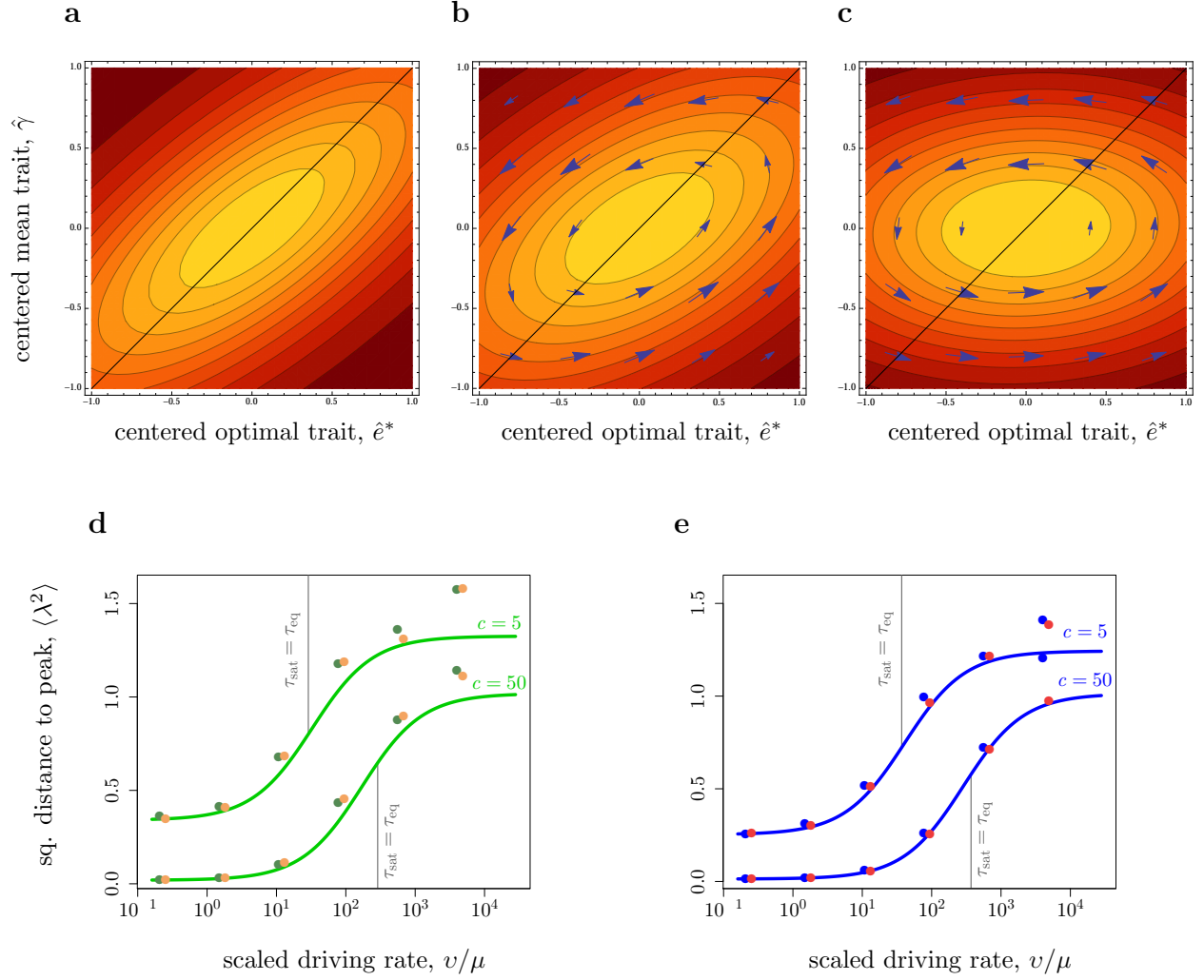


Figure 2: **Stationary distribution of mean and optimal trait in a fitness seascape.** (a–c) The distribution  $Q_{\text{stat}}(\Gamma, E^*)$  is shown (a) in the equilibrium limit ( $c = 1, v = 0, r^2 = 1$ ), (b) for an intermediate driving rate ( $v = 1.5r^2/\tau_{\text{eq}}$ ), and (c) in the deep micro-evolutionary regime ( $v = 50r^2/\tau_{\text{eq}}$ ; see eqs. (28), (36), and (37). The probability current  $\mathbf{J}(\Gamma, E^*)$ , which is given by eq. (35), is marked by arrows. With increasing driving rate, the correlation between  $\Gamma$  and  $E^*$  is seen to decrease. (d,e) The scaled average square distance between mean and optimal trait value,  $\langle \lambda^2 \rangle$ , is plotted against the scaled driving rate  $v/\mu$  for (d) non-recombining and (e) fully recombining populations for different stabilizing strengths  $c$ . The other parameters are  $r^2 = 1, \theta = 0.0125$ . This function increases from an equilibrium value for  $v = 0$  to a micro-evolutionary limit value for  $v \rightarrow \infty$  with a crossover for  $\tau_{\text{sat}}(v, r^2) \sim \tau_{\text{eq}}(c)$ , as given by eq. (33). The analytical results (lines) are compared to simulation results (with parameters  $N = 100, \ell = 100$ ) for a diffusive seascape (green dots) and for a punctuated seascape (orange dots).

where  $\Gamma_a \equiv \Gamma(t_a)$ ,  $\Gamma \equiv \Gamma(t)$ ,  $E_a^* \equiv E^*(t_a)$ ,  $E^* \equiv E^*(t)$ , and  $\tau = t - t_a$ . The function  $G_\tau(\Gamma, E^* | \Gamma_a, E_a^*)$  is the conditional probability (or propagator) for mean and optimal trait values  $\Gamma, E^*$  at time  $t$ , given the values  $\Gamma_a, E_a^*$  at time  $t_a$ .

In a similar way, the average divergence between two descendent populations evolved from a common ancestor population is given by

$$\begin{aligned}
\langle D^{(2)} \rangle(\tau) &\equiv \langle (\Gamma_1(t) - \Gamma_2(t))^2 \rangle_{\substack{\Gamma_1(t_a) = \Gamma_2(t_a) \\ E_1^*(t_a) = E_2^*(t_a)}} \\
&\equiv \int d\Gamma_a d\Gamma_1 d\Gamma_2 (\Gamma_1 - \Gamma_2)^2 \times \left( \begin{array}{c} \text{Diagram: A central point } \Gamma_a \text{ branches into two points } \Gamma_1 \text{ and } \Gamma_2. \text{ Arrows point from } \Gamma_a \text{ to } \Gamma_1 \text{ and } \Gamma_2, \text{ each labeled } \tau/2. \end{array} \right) \\
&= \int dE_a^* dE_1^* dE_2^* d\Gamma_a d\Gamma_1 d\Gamma_2 (\Gamma_1 - \Gamma_2)^2 G_{\tau/2}(\Gamma_1, E_1^* | \Gamma_a, E_a^*) \\
&\quad \times G_{\tau/2}(\Gamma_2, E_2^* | \Gamma_a, E_a^*) Q_{\text{stat}}(\Gamma_a, E_a^*), \tag{41}
\end{aligned}$$

where  $\Gamma_a \equiv \Gamma_1(t_a) = \Gamma_2(t_a)$ ,  $E_a^* \equiv E_1^*(t_a) = E_2^*(t_a)$ ,  $\Gamma_i \equiv \Gamma_i(t)$ ,  $E_i^* \equiv E_i^*(t)$  ( $i = 1, 2$ ), and  $\tau \equiv 2(t - t_a)$ . In Appendix A, we use the Langevin formalism to derive analytical expressions for these propagators and the resulting scaled divergence  $\langle d^{(\kappa)} \rangle(\tau) \equiv \langle D^{(\kappa)} \rangle(\tau) / E_0^2$ , ( $\kappa = 1, 2$ ). We obtain

$$\begin{aligned}
\langle d^{(\kappa)} \rangle(\tau; c, v, r^2) &= \frac{\tau_{\text{eq}}(c)}{\tilde{\tau}(c)} \left[ 1 - e^{-\tau/\tau_{\text{eq}}(c)} \right] \\
&\quad + v w(c, v, r^2) w(c, -v, r^2) \left[ \tau_{\text{sat}}(v, r^2) (1 - e^{-\tau/\tau_{\text{sat}}(v, r^2)}) - \tau_{\text{eq}}(c) (1 - e^{-\tau/\tau_{\text{eq}}(c)}) \right] \\
&\quad - 2(\kappa - 1) \frac{v}{\tau_{\text{eq}}^{-1}(c) + \tau_{\text{sat}}^{-1}(v, r^2)} w(c, -v, r^2)^2 \left[ e^{-\tau/(2\tau_{\text{sat}}(v, r^2))} - e^{-\tau/(2\tau_{\text{eq}}(c))} \right]^2. \tag{42}
\end{aligned}$$

The difference between the two divergence measures is a consequence of the non-equilibrium adaptive dynamics, which violate detailed balance. Equation (42) is valid for diffusive and for punctuated fitness seascapes. It contains the three characteristic time scales defined in the previous section: the *drift time*  $\tilde{\tau}(c)$  is the scale over which the diffusion of the trait mean, in the absence of any fitness seascape, generates a trait divergence of the order of the neutral trait span  $E_0^2$ ; the *equilibration time*  $\tau_{\text{eq}}(c)$  governs the relaxation of the population ensemble to a mutation-selection-drift equilibrium in a fitness landscape of stabilizing strength  $c$ ; the *saturation time*  $\tau_{\text{sat}}$  is defined by the mean square displacement of the fitness peak reaching the driving span  $r^2$ . Here, we focus on fitness seascapes with substantial stabilizing strength and with a driving span of order of the neutral trait span ( $c \gtrsim 1$ ,  $r^2 \sim 1$ ). This selection scenario is biologically relevant: it describes adaptive processes that build up large trait differences by continuous diffusion or recurrent jumps of the fitness peak.

In macro-evolutionary seascapes, the equilibration time and the non-equilibrium saturation time are well-separated,  $\tau_{\text{eq}}(c) \ll \tau_{\text{sat}}(v, r^2)$ . This results in three temporal regimes of the trait divergence:

- **Drift regime**,  $\tau \lesssim \tau_{\text{eq}}(c)$ . The scaled divergence takes the form

$$\begin{aligned}
\langle d^{(\kappa)} \rangle(\tau) &= 2\langle \hat{\gamma}^2 \rangle_{\text{eq}} (1 - e^{-\tau/\tau_{\text{eq}}}) \\
&\simeq \frac{\langle \delta \rangle}{2N} \tau (1 + O(\tau/\tau_{\text{eq}})) \quad (\kappa = 1, 2), \tag{43}
\end{aligned}$$

with an initial increase due to genetic drift and relaxation to an equilibrium value

$$2\langle \hat{\gamma}^2 \rangle_{\text{eq}}(c) = \frac{w(c)}{c} \tag{44}$$

due to stabilizing selection.

- **Adaptive regime**,  $\tau_{\text{eq}}(c) \lesssim \tau \ll \tau_{\text{sat}}(v, r^2)$ . The scaled trait divergence follows

$$\begin{aligned} \langle d^{(\kappa)} \rangle(\tau) &= [2\langle \hat{\gamma}^2 \rangle_{\text{eq}} (1 - v \tilde{\tau} \kappa w(c)^2) + v w(c)^2 \tau] \left[ 1 + O\left(e^{-\tau/\tau_{\text{eq}}}, \tau/\tau_{\text{sat}}\right) \right] \\ &\simeq [2\langle \hat{\gamma}^2 \rangle_{\text{eq}} + v(\tau - \kappa \tau_{\text{eq}}(c))] \left[ 1 + O\left((\theta c)^2, e^{-\tau/\tau_{\text{eq}}}, \tau/\tau_{\text{sat}}\right) \right] \quad (\kappa = 1, 2) \end{aligned} \quad (45)$$

In this regime, the trait divergence is the sum of an (asymptotically constant) equilibrium component and an adaptive component, which increases with slope  $v$ . In a macro-evolutionary fitness seascape, this slope is, by definition, smaller than the slope in the initial drift regime (43), which allows for a clear delineation of the two regimes in empirical data. This feature will be exploited in our selection test for quantitative traits, which will be discussed in section 5.

- **Saturation regime**,  $\tau \gtrsim \tau_{\text{sat}}(v, r^2)$ . On the largest time scales, the divergence

$$\langle d^{(\kappa)} \rangle(\tau) \approx 2\langle \hat{\gamma}^2 \rangle_{\text{eq}} + 2r^2 w(c) w(c, v, r^2) (1 - e^{-\tau/\tau_{\text{sat}}}) \quad (\kappa = 1, 2) \quad (46)$$

approaches its non-equilibrium saturation value

$$\langle \hat{\gamma}^2 \rangle_{\text{stat}}(c, v, r^2) = 2\langle \hat{\gamma}^2 \rangle_{\text{eq}} + 2r^2 w(c) w(c, v, r^2), \quad (47)$$

which equals the  $\Gamma$ -variance of the stationary distribution  $Q_{\text{stat}}(\Gamma, E^*)$ , and is primarily determined by the driving span  $r^2$ . In empirical data, this regime is often well beyond the depth of the phylogeny and, hence, not observable.

In micro-evolutionary seascapes, the saturation of fitness fluctuations occurs faster than the equilibration of the trait under stabilizing selection, i.e.,  $\tau_{\text{sat}}(v, r^2) \lesssim \tau_{\text{eq}}(c)$ . Hence, there is a direct crossover from the drift regime to the saturation regime. For fast micro-evolutionary fitness fluctuations,  $\tau_{\text{sat}}(v, r^2) \ll \tau_{\text{eq}}(c)$ , the constraint on the trait equals that in an effective fitness landscape with stabilizing strength  $c_{\text{eff}}$  given by eq. (39). In this regime, time-dependent trait divergence data alone can no longer resolve adaptive evolution in a fitness seascape from equilibrium in the corresponding effective fitness landscape; this requires additional information on the trait diversity.

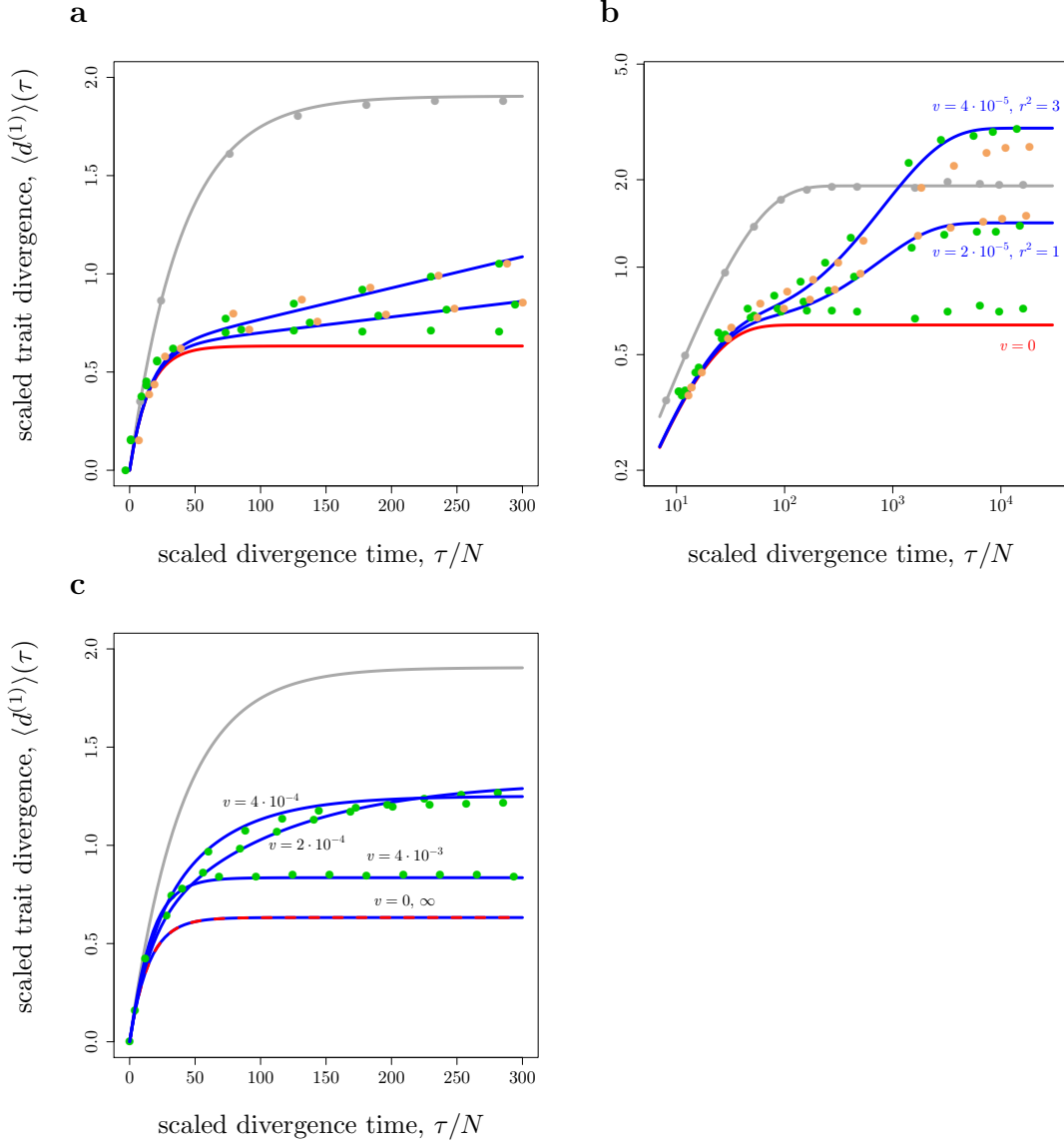
Fig. 3 and Fig. A.1(a–b) show the scaled divergence  $\langle d^{(1)} \rangle(\tau)$  for selection parameters  $c$  and  $v$  covering macro-evolutionary and micro-evolutionary fitness seascapes. The analytical expression of eq. (42) is seen to be in good agreement with numerical simulations for diffusive and punctuated fitness fluctuations.

### 3.3 Stationary trait diversity

As discussed in section 2, our diffusion theory predicts that the movements of the optimum trait in a single-peak fitness seascape of the form (15) only affects the evolution of the trait mean in the population and not the trait diversity. The statistics of the trait diversity remains similar to the case of evolution under stabilizing selection, which is characterized by a time-invariant fitness function,  $F_2(\Delta) = -c_0 \Delta$ . The resulting equilibrium distribution  $Q_{\text{eq}}(\Delta)$  is the product of the neutral mutation-drift equilibrium  $Q_0(\Delta)$ , which is given in eqs. (53) and (55) of [1] and a Boltzmann factor from the scaled fitness landscape,  $Q_{\text{eq}}(\Delta) = Q_0(\Delta) \exp[-c_0 \Delta]$ . These distributions determine the average diversity

$$\langle \Delta \rangle \equiv \int d\Delta \Delta Q_{\text{eq}}(\Delta) \quad (48)$$

and its neutral counterpart  $\langle \Delta \rangle_0$ , as well as the scaled expectation values  $\langle \delta \rangle \equiv \langle \Delta \rangle / E_0^2$  and  $\langle \delta \rangle_0 \equiv \langle \Delta \rangle_0 / E_0^2$ . The selective constraint on the trait diversity enters the diffusion coefficient of the trait mean in eq. (7), which sets the drift time scale  $\tilde{\tau}(c) = (1/2\mu)(\langle \delta \rangle_0 / \langle \delta \rangle(c))$ , as given by eq. (12). The distributions  $Q_0(\Delta)$  and  $Q_{\text{eq}}(\Delta)$  can be written in closed analytical form; unlike in the case of the trait mean, these distributions



**Figure 3: Time-dependence of the trait divergence.** The scaled average divergence  $\langle d^{(1)} \rangle(\tau)$  is shown as a function of the scaled divergence time  $\tau/N$  for three cases: neutral evolution ( $c = 0$ ; grey lines), conservation in a static fitness landscape ( $c = 1, v = 0$ ; red line), and adaptation in a macro-evolutionary fitness seascape ( $c = 1, v > 0$ ; blue lines). Other parameters:  $\theta = 0.0125, N = 100, \ell = 100, \mathcal{E} = 0.7\ell$ . The analytical results of eq. (42) (lines) are compared to simulation results for asexual evolution in diffusive and punctuated fitness landscapes (green and orange dots, respectively). The corresponding results for fully recombining genomes are shown in Fig. A.1. (a,b) Macro-evolutionary landscapes,  $\tau_{sat}(v, r^2) > \tau_{eq}(c)$ . (a) Linear plot: For  $\tau \lesssim \tau_{eq}$ , the trait evolution is dominated by genetic drift and is independent of selection. For  $\tau \gtrsim \tau_{eq}$ , the seascape data show an adaptive divergence component proportional to  $v\tau$ ; the landscape data saturate to an equilibrium divergence set by stabilizing selection. (b) Logarithmic plot: This also shows the non-equilibrium saturation of the seascape data on time scales  $\tau \sim \tau_{sat}$ , when the divergence reaches twice the driving span  $r^2$ . (c) Micro-evolutionary landscapes  $\tau_{sat}(v, r^2) < \tau_{eq}(c)$ . There is a single cross-over from the drift regime for smaller values of  $\tau$  to the saturation regime for larger values of  $\tau$ . The divergence  $\langle d^{(1)} \rangle(\tau)$  equals that in an effective fitness landscape of stabilizing strength  $c_{eff} < c$ . The limit  $v \rightarrow \infty$  has  $c_{eff} = c$ ; i.e., the function  $\langle d^{(1)} \rangle(\tau)$  becomes identical to the case  $v = 0$  (blue-red dashed line).



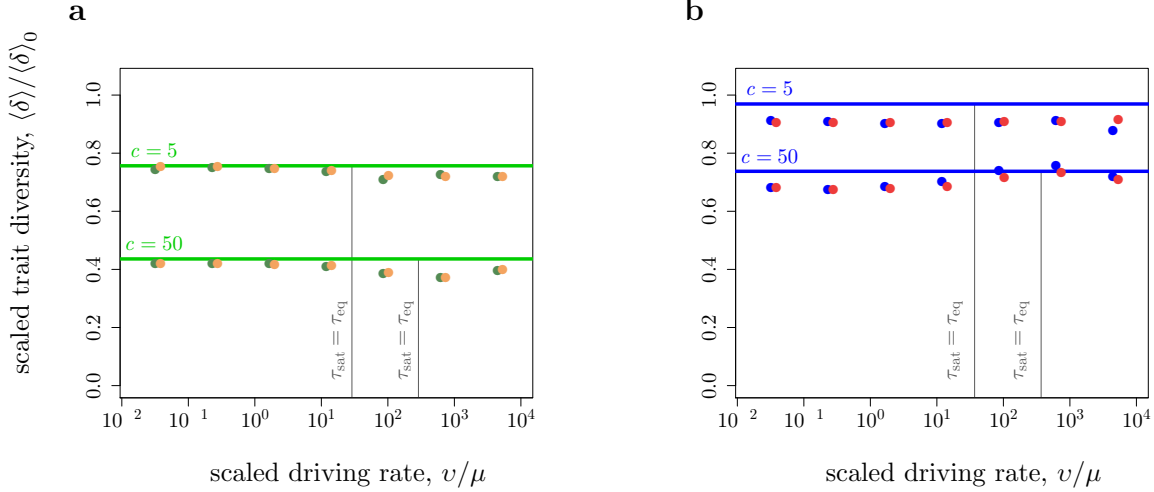


Figure 4: **Stationary trait diversity.** The figure shows the average trait diversity  $\langle \delta \rangle$  (in units of the neutral average  $\langle \delta \rangle_0$ ) in a fitness seascape as a function of the scaled driving rate  $v/\mu$  for different values of the stabilizing strength ( $c = 5, 50$ , top to bottom); other parameters are as in Fig. 3. The equilibrium predictions of diffusion theory (lines), which do not depend on  $v$ , are compared to simulation results of the adaptive process of (a) non-recombining and (b) fully recombining populations in diffusive (green/orange dots) and punctuated (blue/red dots) fitness landscapes. The simulation results confirm evolutionary equilibrium of the trait diversity.

depend directly on the rate of recombination in the population [1]. We obtain the scaled neutral expectation value  $\langle \delta \rangle_0 = 4\theta(1 - 4\theta + O(\theta^2))$ , which is independent of the recombination rate, and the selective constraint

$$\frac{\langle \delta \rangle(c)}{\langle \delta \rangle_0} = \begin{cases} 1 - 4\theta c + O((\theta c)^2) & \text{for } \theta c \ll 1, \\ (4\theta c)^{-1/2} + O((\theta c)^{-1}) & \text{for } \theta c \gg 1 \end{cases} \quad (49)$$

in non-recombining populations. We note that this constraint depends only on the product  $\theta c$ ; therefore, it remains weak over a wide range of parameters ( $c \lesssim 1/\theta$ ), which includes strong selection effects on the trait mean [1]. The full crossover function and the corresponding expressions for fully recombining populations are given in eqs. (68) – (73) of ref. [1].

The numerical simulations reported in Fig. 4 show that the average diversity in diffusive fitness landscapes is well represented by the equilibrium value throughout the crossover from macro- to micro-evolutionary driving rates, and over a wide range of stabilizing strengths. Theoretically, the results of the diffusion theory are valid for adaptive processes unless recurrent selective sweeps reduce the trait diversity within the population. Such sweeps are more prominent in punctuated fitness landscapes due to sudden changes of the trait optimum. We expect a significant reduction in trait diversity due to the large and frequent jumps of the trait optimum in fitness landscapes with very strong stabilizing selection. This regime is beyond the scopes of this paper.

## 4 Fitness and entropy of adaptive processes

The distributions of the trait mean and diversity determine the fitness statistics of an ensemble of populations in the stationary state. These statistics can quantify the cost and the amount of adaption for the evolution of molecular traits. We also evaluate the predictability of the trait evolution in an ensemble of populations after diverging from a common ancestral population.

## 4.1 Genetic load

The genetic load of an individual population is defined as the difference between the maximum fitness and the mean fitness [38–41],

$$L(t) \equiv f^* - \bar{f}(t). \quad (50)$$

For a quantitative trait in a quadratic fitness seascape of the form (15), we can decompose the load into contributions of the trait mean and diversity,

$$L(t) = f^* - c_0(\Gamma(t) - E^*(t))^2 - 2c_0\Delta(t). \quad (51)$$

In the stationary population ensemble (28), the average scaled genetic load can be written as the sum of an equilibrium and an adaptive component,

$$\begin{aligned} \langle 2NL \rangle(c, v, r^2) &= c[\langle \lambda^2 \rangle(c, v, r^2) + \langle \delta \rangle(c)] \\ &= c[\langle \lambda^2 \rangle_{\text{eq}}(c, r^2) + \langle \delta \rangle(c)] + c[\langle \lambda^2 \rangle(c, v, r^2) - \langle \lambda^2 \rangle_{\text{eq}}(c, r^2)] \\ &\equiv 2NL_{\text{eq}}(c, r^2) + 2NL_{\text{ad}}(c, v, r^2); \end{aligned} \quad (52)$$

these components can be computed analytically from eqs. (33) and (34). A simple form is obtained for fitness seascapes of substantial stabilizing strength ( $c \gtrsim 1$ ),

$$2NL_{\text{eq}} \simeq \frac{1}{2} + \mathcal{O}(1/c, \theta c), \quad (53)$$

$$2NL_{\text{ad}}(c, v, r^2) \simeq \begin{cases} v \tilde{\tau}(c) \left[ 1 + \mathcal{O}\left(\frac{\tau_{\text{eq}}}{\tau_{\text{sat}}}\right) \right], & (\text{macroev. seascapes}) \\ cr^2 \left[ 1 - \mathcal{O}\left(\frac{\tau_{\text{sat}}}{\tau_{\text{eq}}}\right) \right], & (\text{microev. seascapes}), \end{cases} \quad (54)$$

where the drift scale  $\tilde{\tau}(c)$  is given by eqs. (12) and (49). From these expressions, we read off three relevant properties of the genetic load.

First, the equilibrium load depends on  $c$  only via its diversity component; this dependence remains weak even for substantial stabilizing selection ( $1 \lesssim c \lesssim 1/\theta$ ). The equilibrium load component related to the trait mean,  $c\langle \lambda^2 \rangle_{\text{eq}}(c)$ , becomes universal in this regime: the fluctuations of  $\Gamma$  are constrained to a fitness range of order  $2NL_{\text{eq}} \simeq 1/2$  around  $E^*$ , irrespectively of the stabilizing strength and the molecular details of the trait [25]. For a  $d$ -component trait as in Fisher’s geometrical model [2], this formula generalizes to  $2NL_{\text{eq}} \simeq d/2$  [42–44]. This is a direct evolutionary analogue of the equipartition theorem in statistical thermodynamics, which states that every degree of freedom that enters the energy function quadratically contributes an average of  $k_B T/2$  to the total energy of a system at temperature  $T$  (the proportionality factor  $k_B$  is Boltzmann’s constant). [25].

Similarly, the adaptive load component depends only weakly on  $c$ , via the drift scale  $\tilde{\tau}(c)$ . At a fixed value of  $\Gamma$ , the stochastic displacement of the fitness peak induces a fitness cost proportional to  $c$ ; however, this effect is largely offset by an adaptive response that becomes faster with increasing  $c$ .

Finally, the different regimes of adaptive trait evolution can be characterized in terms of the genetic load. The adaptive load is asymptotically linear in the driving rate and is subleading to the equilibrium load in the slow-driving regime ( $v \lesssim \tilde{v}(c) \equiv 1/\tilde{\tau}(c)$ ). It becomes dominant for faster driving ( $v \gtrsim \tilde{v}(c)$ ) and saturates in the micro-evolutionary regime ( $v \gtrsim r^2/\tau_{\text{eq}}(c)$ ). Fig. 5(a) shows this dependence of the adaptive load on the driving rate.

## 4.2 Fitness flux

The fitness flux,  $\phi(t)$ , characterizes the adaptive response of a population evolving in a fitness land- or seascape,

$$\phi(t) = \int dE f(E, t) \frac{\partial}{\partial t} \mathcal{W}(E, t). \quad (55)$$

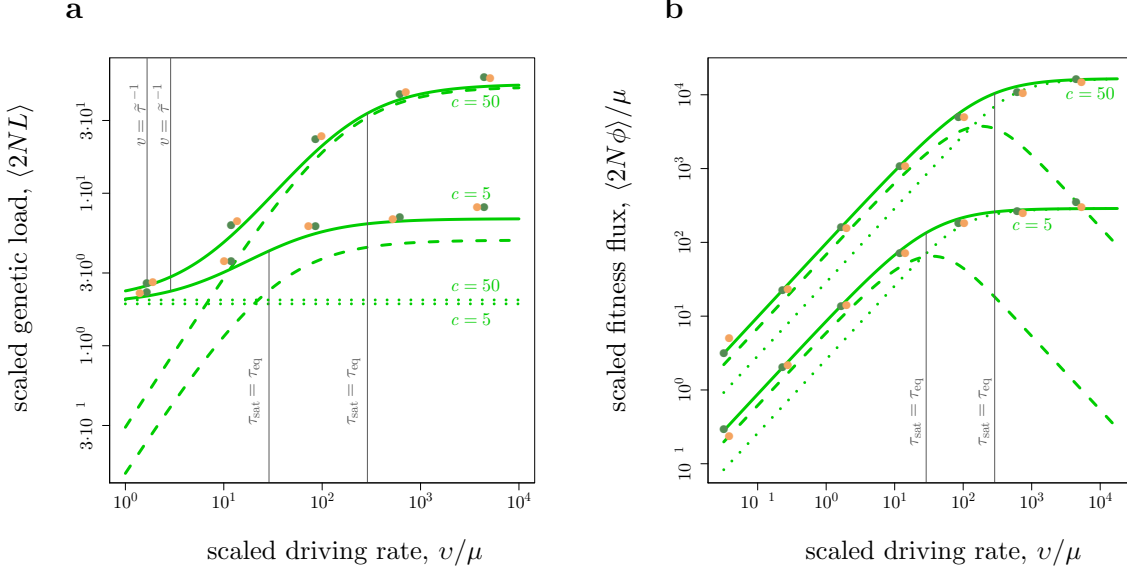


Figure 5: **Genetic load and fitness flux.** (a) Scaled genetic load  $2NL$  (full lines) and its constituents, the adaptive genetic load,  $2NL_{ad}$  (dashed lines), and equilibrium genetic load,  $2NL_{eq}$  (dotted lines), for stationary evolution of non-recombining populations in a fitness seascape. The load components are plotted against the scaled driving rate  $v/\mu$  for stabilizing strengths  $c = 5, 50$ ; other parameters like in Fig. 3. The analytical results of eq. (52) are compared to simulations for diffusive and punctuated fitness landscapes (green and orange dots). The corresponding data for fully recombining populations are shown in Fig. A.1. The genetic load is dominated for  $v \lesssim 1/\tilde{\tau}(c)$  by the equilibrium component and for  $v \gtrsim 1/\tilde{\tau}(c)$  by the adaptive component; it saturates in the micro-evolutionary seascape regime ( $v \gtrsim r^2/\tau_{eq}(c)$ ). (b) The scaled fitness flux  $\langle 2N\phi \rangle$  (solid line) and its components  $\langle 2N\phi_{macro} \rangle$  and  $\langle 2N\phi_{micro} \rangle$ , as defined in eqs. (63), are shown for the same parameters (all flux values are measured in units of  $1/\mu$ ). In macro-evolutionary fitness landscapes,  $\langle 2N\phi \rangle$  is an approximately linear function of the driving rate  $v$  and the component  $\langle 2N\phi_{macro} \rangle$  is the dominant part. In micro-evolutionary landscapes,  $\langle 2N\phi \rangle$  saturates and the component  $\langle 2N\phi_{micro} \rangle$  is the dominant part.

The cumulative fitness flux,  $\Phi(\tau) = \int_t^{t+\tau} \phi(t') dt'$ , measures the total amount of adaptation over an evolutionary period  $\tau$  [45]. The evolutionary statistics of this quantity is specified by the fitness flux theorem [17]. According to the theorem, the average cumulative fitness flux in a population ensemble measures the deviation of the evolutionary process from equilibrium: this deviation equals the relative entropy of the actual process from a hypothetical time-reversed process [17, 46]. It is substantial — i.e., the process is predominantly adaptive — if  $\langle 2N\Phi \rangle \gtrsim 1$ . Specifically, the cumulative fitness flux of a stationary adaptive process increases linearly with time,  $\langle 2N\Phi(\tau) \rangle = \langle \phi \rangle \tau$  with  $\langle \phi \rangle > 0$ .

For a quantitative trait in a quadratic fitness seascape of the form (15), we can decompose the fitness flux into contributions of the trait mean and the trait diversity,

$$\phi(t) = -2c_0(\Gamma(t) - E^*(t)) \frac{d\Gamma(t)}{dt} - 2c_0 \frac{d\Delta(t)}{dt}. \quad (56)$$

In the stationary population ensemble (28), the average scaled fitness flux can be expressed in terms of the stationary probability current  $\mathbf{J}_{stat}(\Gamma, E^*)$ ,

$$\langle 2N\phi \rangle = -\frac{2c}{E_0^2} \int d\Gamma dE^* (\Gamma - E^*) J_{stat}^\Gamma(\Gamma, E^*), \quad (57)$$

where  $J_{stat}^\Gamma(\Gamma, E^*)$  is the  $\Gamma$ -component of  $\mathbf{J}_{stat}(\Gamma, E^*)$ . The fitness flux can be computed analytically from eq. (35),

$$\langle 2N\phi \rangle(c, v, r^2) = 2cv w(c, v, r^2). \quad (58)$$

In the regime of substantial stabilizing strength ( $c \gtrsim 1$ ), we get

$$\langle 2N\phi \rangle(c, v, r^2) \simeq \begin{cases} 2cv \left[ 1 - \mathcal{O}\left(\frac{\tau_{\text{eq}}}{\tau_{\text{sat}}}\right) \right] & (\text{macroev. seascapes}), \\ \frac{4c^2 r^2}{\tilde{\tau}(c)} \left[ 1 - \mathcal{O}\left(\frac{\tau_{\text{sat}}}{\tau_{\text{eq}}}\right) \right] & (\text{microev. seascapes}), \end{cases} \quad (59)$$

where the drift time  $\tilde{\tau}(c)$  is given by eqs. (12) and (49). The fitness flux depends linearly on the driving rate in a macro-evolutionary fitness seascape, and it saturates in the regime of micro-evolutionary fitness fluctuations. Fig. 5(b) shows this dependence of the fitness flux on the driving rate.

We can express the fitness flux in terms of correlation functions of the trait mean  $\Gamma(t)$  and the lag  $\Lambda(t)$ , which results in a simple relation between fitness flux and adaptive load. Inserting the probability current of eq. (35) into the integral of eq. (57), we find

$$\begin{aligned} \langle 2N\phi \rangle &= \frac{2c^2 \langle \delta \rangle}{E_0^2} (\langle \Lambda^2 \rangle - \langle \Lambda^2 \rangle_{\text{eq}}) + \frac{4c\theta}{E_0^2} \lim_{\tau \searrow 0} (\langle \Lambda(t+\tau)(\Gamma(t) - \Gamma_0) \rangle - \langle \Lambda(t+\tau)(\Gamma(t) - \Gamma_0) \rangle_{\text{eq}}) \\ &= c \langle \delta \rangle \langle 2NL \rangle_{\text{ad}}(c, v, r^2) [1 + \mathcal{O}(\theta)]. \end{aligned} \quad (60)$$

From this representation, we obtain the spectral decomposition of the fitness flux,

$$\langle 2N\phi \rangle(c, v, r^2) = \int_0^\infty \langle 2N\phi(\omega) \rangle d\omega \quad (61)$$

with

$$\langle 2N\phi(\omega) \rangle = 2cv \frac{c \langle \delta \rangle}{\pi/2} \frac{\omega^2}{(\tau_{\text{eq}}^{-2}(c) + \omega^2)(\tau_{\text{sat}}^{-2}(v, r^2) + \omega^2)} [1 + \mathcal{O}(\theta/(c \langle \delta \rangle))]. \quad (62)$$

Using a cutoff frequency  $\omega_c = k/\tau_{\text{eq}}(c)$  with a constant  $k$  of order 1, we can now define a macro-evolutionary flux component,

$$\begin{aligned} \langle 2N\phi_{\text{macro}} \rangle &= \int_0^{\omega_c} \langle 2N\phi(\omega) \rangle d\omega \\ &= 2cv w(c, v, r^2) \frac{(\tau_{\text{eq}}^{-1}(c) + 2\mu) \arctan[k] - (\tau_{\text{sat}}^{-1}(v, r^2) + 2\mu) \arctan[k \tau_{\text{sat}}(v, r^2)/\tau_{\text{eq}}(c)]}{(\pi/2)(\tau_{\text{eq}}^{-1}(c) - \tau_{\text{sat}}^{-1}(v, r^2))}, \end{aligned} \quad (63)$$

and the complementary micro-evolutionary component

$$\langle 2N\phi_{\text{micro}} \rangle = \int_{\omega_c}^\infty \langle 2N\phi(\omega) \rangle d\omega = \langle 2N\phi \rangle - \langle 2N\phi_{\text{macro}} \rangle. \quad (64)$$

In the regime of substantial stabilizing selection ( $c \gtrsim 1$ ), the macro-evolutionary fitness flux in (63) reads

$$\langle 2N\phi_{\text{macro}} \rangle(c, v, r^2) \simeq \begin{cases} 2cv \frac{2}{\pi} \arctan[k], & (\text{macroev. seascapes}) \\ 2cv \frac{\tau_{\text{sat}}^2(v, r^2)}{\tau_{\text{eq}}^2(c)} \frac{2}{\pi} (k - \arctan[k]) \sim \frac{1}{v} & (\text{microev. seascapes}). \end{cases} \quad (65)$$

This fitness flux component quantifies the macro-evolutionary part of adaptation. In macro-evolutionary fitness seascapes ( $\tau_{\text{sat}}(v, r^2) \gtrsim \tau_{\text{eq}}(c)$ ), it increases proportionally to the driving rate  $v$  and, for  $k > 1$ , it represents the main fraction of the total fitness flux  $\langle 2N\phi \rangle$ . In micro-evolutionary fitness seascapes ( $\tau_{\text{sat}}(v, r^2) \lesssim \tau_{\text{eq}}(c)$ ), this component is suppressed: the macro-evolutionary fitness flux does not carry

information on rapid fitness fluctuations. This cross-over of  $\langle 2N\phi_{\text{macro}} \rangle$  and of the complementary component  $\langle 2N\phi_{\text{micro}} \rangle$  is shown in Fig. 5(b).

The spectral decomposition of the fitness flux has important consequences for the analysis of macro-evolutionary adaptation. The detection of a substantial cumulative fitness flux  $\langle 2N\Phi_{\text{macro}}(\tau) \rangle > 1$  over a macro-evolutionary period  $\tau$  is not confounded by the simultaneous presence of micro-evolutionary (for example seasonal) fitness fluctuations. Since the cumulative fitness flux is a measure of entropy production during adaptation, the spectral decomposition (61) also has an important information-theoretic interpretation: The difference  $\langle 2N\phi_{\text{micro}} \rangle = \langle 2N\phi \rangle - \langle 2N\phi_{\text{macro}} \rangle$  is the average loss of information per unit time through temporal coarse-graining. This loss is a non-equilibrium analogue of the entropy production by spatial coarse-graining.

### 4.3 Predictability and entropy production

In ref. [1], we quantified the evolutionary predictability of the molecular traits across an ensemble of populations by

$$\mathcal{P} \equiv \exp [\langle S(\mathcal{W}) \rangle - S(\langle \mathcal{W} \rangle)], \quad (66)$$

with  $S(\mathcal{W}) \equiv -\int \mathcal{W}(E) \log \mathcal{W}(E) dE$ . This definition compares the ensemble-averaged “micro-evolutionary” Shannon entropy of the phenotype distribution within a population,  $\langle S(\mathcal{W}) \rangle \equiv \int_{\mathcal{W}} S(\mathcal{W}) Q(\mathcal{W})$ , and the “macro-evolutionary” Shannon entropy of the mixed distribution,  $S(\langle \mathcal{W} \rangle) \equiv S(\int_{\mathcal{W}} \mathcal{W} Q(\mathcal{W}))$ , which is obtained by compounding the trait values of all populations into a single distribution. We have shown that the predictability is generically low in a neutral ensemble, but stabilizing selection in a single fitness landscape can generate an evolutionary equilibrium with predictability values  $\mathcal{P}$  of order 1 [1].

Here we compute the predictability in a time-dependent ensemble of populations that descend from a common ancestor population. Similarly to ref. [1], we evaluate eq. (66) for a distribution  $Q_t(\mathcal{W})$  with the initial condition  $Q_{t_a}(\mathcal{W}) = \delta(\mathcal{W} - \mathcal{W}_a)$  at time  $t_a = t - \tau/2$ . We obtain the time-dependent predictability

$$\mathcal{P}(\tau; c, v, r^2) \simeq \left( \frac{\langle \delta \rangle(c)}{\langle d^{(2)} \rangle(\tau; c, v, r^2)/2 + \langle \delta \rangle(c)} \right)^{1/2} = \left( \frac{1}{1 + \Omega^{(2)}(\tau; c, v, r^2)/4\theta} \right)^{1/2}. \quad (67)$$

Here,  $\Omega^{(2)}(\tau; c, v, r^2) \equiv 2\theta \langle d^{(2)} \rangle(\tau; c, v, r^2)/\langle \delta \rangle$  denotes the ratio between trait divergence and diversity for the descendent populations. The trait statistics in a macro-evolutionary fitness seascape, given by eqs. (42) and (49), entail the evolutionary predictability

$$\mathcal{P}(\tau; c, v, r^2) = \mathcal{P}_{\text{eq}}(c) \left[ 1 - \frac{1}{2} v \tilde{\tau} \frac{\tau - 2\tau_{\text{eq}}(c)}{2N} \left[ 1 + \mathcal{O}(\tau \tau_{\text{eq}}(c) cv/N, \tau/\tau_{\text{sat}}, \theta/(c\langle \delta \rangle)) \right] \right] \quad (68)$$

for  $\tau \gtrsim \tau_{\text{eq}}(c)$ , with  $\mathcal{P}_{\text{eq}}(c) = (1+w(c)/(2c))^{-1/2} = (1+1/(2c))^{-1/2} [1 + \mathcal{O}(\theta/(c\langle \delta \rangle))]$ . There are two stochastic components that generate macro-evolutionary entropy and, hence, reduce the evolutionary predictability: fluctuations induced by genetic drift on short time-scales  $\tau \lesssim \tau_{\text{eq}}(c)$  and fluctuations of the fitness peak over time-scales  $\tau \gtrsim \tau_{\text{eq}}(c)$ . Stabilizing selection, on the other hand, reduces the entropy production of the adaptive process [25]. Therefore, the predictability of an adaptive process in a fitness seascape with a substantial stabilizing strength and sufficiently slow driving rate can remain of order one.

Parallel and convergent evolution at the functional level, paired with strongly divergent genome evolution has been observed in a number of recent experiments [47–49]. These experimental observations can be explained in a natural way, if we assume that many of these functions involve a complex quantitative trait.

## 5 Inference of adaptive trait evolution

The statistical theory developed in this paper suggests a new method to infer selection on quantitative traits. Our method is based on trait evolution in a single-peak fitness seascape, as defined in eq. (15), which is parametrized by its stabilizing strength  $c$  and its driving rate  $v$ .

Two main results are relevant for the inference of selection. First, evolution in a macro-evolutionary fitness seascape affects the population mean trait in complementary ways: it generates conservation on shorter scales and adaptation on longer scales of evolutionary time. These characteristics are measured by the expected trait divergence between populations,  $\langle D^{(\kappa)}(\tau) \rangle$ , which depends on the divergence time  $\tau$  and on the selection parameters  $c$  and  $v$  in a characteristic way. The divergence can be measured either between an ancestral population and a descendent population ( $\kappa = 1$ ) or between two descendent populations evolving from a common ancestor population ( $\kappa = 2$ ). As discussed in section 3.2, these measures are generically distinct for adaptive processes<sup>2</sup>. Second, the expected trait diversity within populations,  $\langle \Delta \rangle$ , shows a weaker signal of conservation. Moreover, it decouples from the adaptive process in a single-peak fitness seascape over a wide range of evolutionary parameters, as discussed in section 3.3.

Our test statistics is the time-dependent divergence-diversity ratio

$$\Omega^{(\kappa)}(\tau) = 2\theta \frac{\langle D^{(\kappa)}(\tau) \rangle}{\langle \Delta \rangle} \quad (\kappa = 1, 2), \quad (69)$$

where  $\theta = \mu N$  denotes the nucleotide diversity. This function depends on the divergence time  $\tau$  and on the selection parameters  $c$  and  $v$ . The typical behavior of  $\Omega^{(\kappa)}(\tau)$  for different evolutionary modes is shown in Fig. 6 and can be summarized as follows:

- *Neutral evolution* ( $c = 0$ ). The divergence-diversity ratio has an initially linear increase due to genetic drift and approaches a maximum value 1 with a relaxation time  $\tau_0 = 1/\mu$ ,

$$\Omega^{(\kappa)}(\tau) = \Omega_0(\tau) \simeq \begin{cases} \mu\tau & \text{for } \tau \ll \tau_0 \\ 1 & \text{for } \tau \gg \tau_0 \end{cases} \quad (\kappa = 1, 2). \quad (70)$$

The function  $\Omega_0(\tau)$ , which does not depend on  $\kappa$  by detailed balance, is shown as a grey line in Fig. 6.

- *Conservation in a fitness landscape* ( $c \gtrsim 1, v = 0$ ). The divergence-diversity ratio approaches a smaller maximum value,  $\Omega_{\text{stab}}(c) < 1$ , with a proportionally shorter relaxation time  $\tau_{\text{eq}}(c) = \Omega_{\text{stab}}(c)/\mu$ ,

$$\Omega^{(\kappa)}(\tau) = \Omega_{\text{eq}}(\tau; c) \simeq \begin{cases} \mu\tau & \text{for } \tau \ll \tau_{\text{eq}}(c) \\ \Omega_{\text{stab}}(c) & \text{for } \tau \gg \tau_{\text{eq}}(c) \end{cases} \quad (\kappa = 1, 2). \quad (71)$$

The function  $\Omega_{\text{eq}}(\tau; c)$ , which does not depend on  $\kappa$  by detailed balance, is shown as a red line in Fig. 6. Over a wide range of evolutionary parameters, the maximum value depends on the stabilizing strength in a simple way,  $\Omega_{\text{stab}}(c) \sim 1/(2c)$ , with corrections for weaker selection and for larger nucleotide diversity.

- *Adaptation in a macro-evolutionary fitness seascape* ( $c \gtrsim 1, 0 < v \lesssim 1/\tilde{\tau}$ ). The divergence-diversity ratio acquires an adaptive component,

$$\begin{aligned} \Omega^{(\kappa)}(\tau) &= \Omega_{\text{eq}}(\tau; c) + \Omega_{\text{ad}}^{(\kappa)}(\tau; v) \\ &= \Omega_{\text{eq}}(\tau; c) + \frac{v}{2} [\tau - \kappa \tau_{\text{eq}}(c)] \quad (\kappa = 1, 2), \end{aligned} \quad (72)$$

with corrections for weaker selection and for  $\tau$  approaching the non-equilibrium saturation time  $\tau_{\text{sat}} = r^2/v$ . The functions  $\Omega^{(\kappa)}(\tau)$  are shown as blue lines in Fig. 6.

The  $\Omega$ -test for selection on quantitative traits is conceptually related to the McDonald-Kreitman test for adaptive sequence evolution [50]; both tests are based on a comparison between divergence and diversity. However, the  $\Omega$  statistics does not require a corresponding “null trait” that evolves near neutrality and takes the role of synonymous sequences. Indeed, no such neutral trait gauge is available in most cases. The  $\Omega$ -test

<sup>2</sup>The relative difference between  $\langle D^{(1)}(\tau) \rangle$  and  $\langle D^{(2)}(\tau) \rangle$  is small (Fig. 5). This difference is conceptually important, however, because it manifests the violation of detailed balance in adaptive processes. Similar effects are ubiquitous in divergence data of trait adaptation across multi-branch phylogenies.

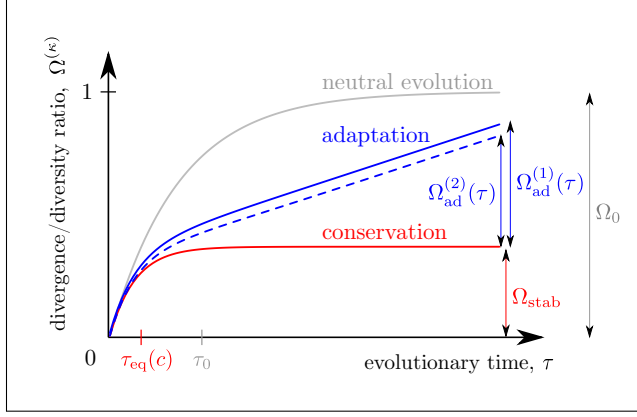


Figure 6: **The universal divergence-diversity ratio**  $\Omega^{(\kappa)}$  ( $\kappa = 1, 2$ ), as defined in eq. (69), for a quantitative trait evolving in a single-peak fitness land- or seascape. This ratio is plotted as a function of the scaled divergence time,  $\tau$ . *Neutral evolution*: The function  $\Omega_0(\tau)$  is independent of  $\kappa$  and reaches its saturation value 1 on times scales  $\tau \gg \tau_0 = 1/\mu$  (grey curve). *Conservation in a fitness landscape*: The function  $\Omega_{\text{eq}}(\tau)$  is independent of  $\kappa$  and has a smaller saturation value  $\Omega_{\text{stab}}(c)$  reached faster than for neutral evolution, on time scales  $\tau \gg \tau_{\text{eq}}(c)$  (red curve). *Adaptation in a fitness seascape*: There is a linear surplus  $\Omega_{\text{ad}}^{(\kappa)}(\tau) \simeq v[\tau - \kappa\tau_{\text{eq}}(c)]$ , which measures the amount of adaptation (blue curves).

instead evaluates *time-resolved* divergence data  $D(\tau)$ . In macro-evolutionary fitness seascapes, this test infers stabilizing and directional selection using their different characteristic time scales. In principle, a single data point in the saturation regime  $\tau \gtrsim \tau_{\text{eq}}(c)$  determines the stabilizing strength of a fitness landscape, and two data points are sufficient to determine strength and driving rate of a fitness seascape. Our statistical theory also specifies an error model for the probabilistic inference of seascape parameters from noisy data, based on the statistics of the finite time propagator detailed in Appendix A. This inference method can be generalized from single lineages to multi-species phylogenies and will be discussed in a forthcoming empirical study.

An important prerequisite for the wide applicability of the  $\Omega$  test is its *universality*: the divergence-diversity ratio depends on the selection parameters  $c$  and  $v$ , but it decouples from the trait’s genetic basis. In particular, it depends only weakly on the number and trait amplitudes of the constitutive sequence sites, and on the amount of recombination between these sites. All of these genetic factors are, in general, unknown. They act as confounding factors on non-universal observables such as the trait divergence and diversity, which have been used in previous studies to infer selection [16, 18]. The  $\Omega$  statistics also decouples from details of the selection dynamics; it can be applied to continual as well as to punctuated adaptive processes. We have tested this universality by extensive numerical simulations, which are reported in Appendix B.

## 6 Discussion

In this paper, we have developed a statistical theory for the evolution of a quantitative trait in a stochastic fitness seascape. The fitness model used for our analysis, a single-peak seascape with diffusive or punctuated peak displacements, covers a broad spectrum of biologically relevant evolutionary scenarios [25]. The two seascape parameters  $c$  and  $v$  quantify stabilizing and directional selection on the trait, which, in turn, govern the trait’s fundamental evolutionary modes of conservation and adaptation. Our analysis shows that these modes are not mutually exclusive, but are joint features of dynamic selection models.

In a macro-evolutionary fitness seascape, conservation and adaptation are associated with different time scales: conservation is observed on shorter scales, while adaptive changes build up on longer scales of evolutionary time. Micro-evolutionary fitness fluctuations, on the other hand, lead to reduced genetic adaptation, which decouples from the macro-evolutionary dynamics of the trait. Rapid adaptive response to seasonal or other fluctuations of the environment often involves epigenetic modifications or phenotypic switching [36].

The evolutionary roles of these mechanisms are beyond the scope of this paper. The spectral decomposition of the fitness flux, which has been introduced above, quantifies how the adaptive process is distributed on different scales of evolutionary time.

Our theory suggests a new test for selection on quantitative traits, which has important potential applications. At the sequence level, an increasingly complex picture of selection has emerged in recent years. Notably, we have acquired a growing repertoire of empirical genotype-fitness landscapes [51], which has generated important experimental and theoretical insights into the evolutionary dynamics on these landscapes. However, we still know little about the statistical properties of empirical phenotype-fitness maps, and next to nothing about phenotype-dependent seascapes. Systematic inference of selection on molecular quantitative traits, such as levels of gene expression and enzymatic activity, can contribute to close this gap. Eventually, fitness land- and seascapes for individual traits will need to be integrated into larger phenotype-fitness maps, which include fitness interactions between traits.

## References

- [1] A. Nourmohammad, S. Schiffels, and M. Lässig. Evolution of molecular phenotypes under stabilizing selection. *J. Stat. Mech.*, P01012, 2013.
- [2] R. Fisher. *The genetical theory of natural selection*. Oxford University Press, USA, 1st edition, 1930.
- [3] R. Lande. Natural-selection and random genetic drift in phenotypic evolution. *Evolution*, 30(2):314–334, 1976.
- [4] N. H. Barton and M. Turelli. Evolutionary quantitative genetics: how little do we know? *Annu Rev Genet*, 23:337–370, 1989.
- [5] D. S. Falconer. *Introduction to quantitative genetics*. Halsted Press, 1989.
- [6] M. Lynch and B. Walsh. *Genetics and analysis of quantitative traits*. Sinauer Associates Inc, 1998.
- [7] S. Rice. A geometric model for the evolution of development. *J Theor Biol*, 1990.
- [8] D. L. Hartl and C. H. Taubes. Compensatory nearly neutral mutations: selection without adaptation. *J Theor Biol*, 182(3):303–309, 1996.
- [9] M. G. Blumer. The genetic variability of polygenic characters under optimizing selection, mutation and drift. *Genet Res*, 19(1):17–25, 1972.
- [10] N. H. Barton. The maintenance of polygenic variation through a balance between mutation and stabilizing selection. *Genet Res*, 47(3):209–216, 1986.
- [11] S. Wright. The analysis of variance and the correlations between relatives with respect to deviations from an optimum. *J Genet*, 30(2):243–256, 1935.
- [12] V. Mustonen and M. Lässig. From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. *Trends Genet*, 25(3):111–9, 2009.
- [13] M. Kopp and J. Hermisson. The genetic basis of phenotypic adaptation ii: the distribution of adaptive substitutions in the moving optimum model. *Genetics*, 183:1453–76, 2009.
- [14] H. P. de Vladar and N. H. Barton. The statistical mechanics of a polygenic character under stabilizing selection, mutation and drift. *J R Soc Interface*, 8(58):720–739, 2011.
- [15] V. Mustonen and M. Lässig. Molecular evolution under fitness fluctuations. *Phys Rev Lett*, 100:108101, 2008.



- [16] T. Bedford, D. L Hartl. Optimization of gene expression by natural selection. *Proc Natl Acad Sci USA* 106: 1133–1138, 2009.
- [17] V. Mustonen and M. Lässig. Fitness flux and ubiquity of adaptive evolution. *Proc Natl Acad Sci*, 107:4248–53, 2010.
- [18] Y. Gilad, A. Oshlack, G. K. Smyth, T. P. Speed, and K. P. White. Expression profiling in primates reveals a rapid evolution of human transcription factors. *Nature* 440: 242–245, 2006.
- [19] M. Kimura. Diffusion models in population genetics. *J Appl Probab*, 1:177–232, 1964.
- [20] W. J. Ewens. *Mathematical Population Genetics*. Springer, New York, 2004.
- [21] R. A. Neher and B. I. Shraiman. Competition between recombination and epistasis can cause a transition from allele to genotype selection. *Proc Natl Acad Sci*, 106(16):6866–6871, 2009.
- [22] R. A. Neher, M. Vucelja, M. Mézard and B. I. Shraiman. Emergence of clones in sexual populations. *J Stat Mech*, P01008, 2013.
- [23] J. Berg, S. Willmann, and M. Lässig. Adaptive evolution of transcription factor binding sites. *BMC Evol Biol*, 4(1):42, 2004.
- [24] M. Lässig. From biophysics to evolutionary genetics: statistical aspects of gene regulation. *BMC Bioinformatics* 8: S7, 2007.
- [25] A. Nourmohammad, T. Held, and M. Lässig. Universality and predictability in molecular quantitative genetics. *Curr Opin Genet Dev*, 23(6):684 – 693, 2013. Genetics of system biology.
- [26] V. Mustonen and M. Lässig. Evolutionary population genetics of promoters: predicting binding sites and functional phylogenies. *Proc Natl Acad Sci*, 102(44):15936–15941, 2005.
- [27] N. H. Barton and J. B. Coe. On the application of statistical physics to evolutionary biology. *J Theor Biol*, 259(2):317–324, 2009.
- [28] N. H. Barton and H. P. de Vladar. Statistical Mechanics and the Evolution of Polygenic Quantitative Traits. *Genetics*, 181(3):997–1011, 2009.
- [29] R. A. Neher and B. I. Shraiman. Statistical genetics and evolution of quantitative traits. *Rev Mod Phys*, 83(4):1283–1300, 2011.
- [30] F. Poelwijk, P. Heyning, M. de Vos, D. Kiviet, and S. Tans. Optimality and evolution of transcriptionally regulated gene expression. *BMC Systems Biology*, 5(1):128, 2011.
- [31] U. Gerland and T. Hwa. On the selection and evolution of regulatory DNA motifs. *J Mol Evol*, 55(4):386–400, 2002.
- [32] V. Mustonen, J. Kinney, C. G. J. Callan, and M. Lässig. Energy-dependent fitness: A quantitative model for the evolution of yeast transcription factor binding sites. *Proc Natl Acad Sci*, 105(34):12376–12381, 2008.
- [33] C. S. Wylie and E. I. Shakhnovich. A biophysical protein folding model accounts for most mutational fitness effects in viruses. *Proc Natl Acad Sci*, 108(24):9916–9921, 2011.
- [34] R. Hermesen, J. B. Deris, and T. Hwa. On the rapidity of antibiotic resistance evolution facilitated by a concentration gradient. *Proc Natl Acad Sci*, 2012.
- [35] M. Lässig. Dynamical anomalies and intermittency in burgers turbulence. *Phys Rev Lett*, 84(12):2618–21, 2000.

- [36] O. Rivoire and S. Leibler. A model for the generation and transmission of variations in evolution. *arXiv:1312.0226*, preprint, 2013.
- [37] M. Möhle. Forward and backward diffusion approximations for haploid exchangeable population models. *Stoch Proc Appl*, 95:133–149, 2001.
- [38] J. F. Crow. Some possibilities for measuring selection intensities in man. *Hum Biol*, 30(1):1–13, 1958.
- [39] J. F. Crow and M. Kimura. Evolution in sexual and asexual populations. *Am Nat*, 99(909):439–450, 1965.
- [40] J. B. S. Haldane. The cost of natural selection. *J Genet*, 55(3):511–524, 1957.
- [41] H. J. Muller. Our load of mutations. *Am J Hum Genet*, 2(2):111–176, 1950.
- [42] D. L. Hartl and C. H. Taubes. Towards a theory of evolutionary adaptation. *Genetica*, 102-103:525–533, 1998.
- [43] A. Poon and S. P. Otto. Compensating for our load of mutations: freezing the meltdown of small populations. *Evolution*, 54:1467–1479, 2000.
- [44] O. Tenaillon, O. K. Silander, J.-P. Uzan, and L. Chao. Quantifying organismal complexity using a population genetic approach. *PloS one*, 2(2):e217, 2007.
- [45] V. Mustonen and M. Lässig. Adaptations to fluctuating selection in *Drosophila*. *Proc. Natl. Acad. Sci*, 104:2277–82, 2007.
- [46] V. Y. Chernyak, M. Chertkov, and C. Jarzynski. Path-integral analysis of fluctuation theorems for general langevin processes. *J Stat Mech*, 8:080001, 2006.
- [47] O. Tenaillon, A. Rodríguez-Verdugo, R. L. Gaut, P. McDonald, A. F. Bennett, A. D. Long, and B. S. Gaut. The molecular diversity of adaptive convergence. *Science*, 335:457–461, 2012.
- [48] E. Toprak, A. Veres, J. B. Michel, R. Chait, D. L. Hartl, and R. Kishony. Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. *Nat. Genet.*, 44:101–105, 2012.
- [49] J. Barroso-Batista, A. Sousa, M. Lourenco, M. Bergman, J. Demengeot, and I. Gordo. The first steps of adaptation of *escherichia coli* to the gut are dominated by soft sweeps. *arXiv:1311.2435*, 2013.
- [50] J. H. McDonald and M. Kreitman. Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature*, 351(6328):652–654, 1991.
- [51] I. G. Szendro, J. Franke, J. A. G. M. de Visser, and J. Krug. Predictability of evolution depends nonmonotonically on population size. *Proc Natl Acad Sci*, 110(2):571–576, 2013.
- [52] C. Gardiner. *Handbook of Stochastic methods: for physics, chemistry and the natural sciences*. Springer, 3rd edition, 2004.

## A Analytical theory of the adaptive ensemble

In Section 3.1, we obtained the Gaussian stationary distribution  $Q_{\text{stat}}(\Gamma, E^*)$  in a diffusive seascape from the underlying Fokker-Planck equation (7). Here we use a Langevin representation to compute the time-resolved trait divergence  $\langle d^{(\kappa)} \rangle(\tau)$  ( $\kappa = 1, 2$ ). This derivation, which reproduces mean and variance of the distribution  $Q_{\text{stat}}(\Gamma, E^*)$ , applies to diffusive and punctuated fitness seascapes. We also compute the full propagator function  $G_\tau(\Gamma, E^* | \Gamma_a, E_a^*)$  for macro-evolutionary diffusive seascapes. The propagator in a punctuated fitness seascape has the same mean and variance, but differs in higher trait moments.

**Moments of the optimal trait.** In a diffusive seascape, the fitness peak  $E^*(t)$  follows an Ornstein-Uhlenbeck process with Langevin representation

$$\partial_t E^*(t) = -\frac{v}{r^2}(E^*(t) - \mathcal{E}) + \eta(t), \quad (73)$$

where  $\eta(t)$  is a Gaussian random variable with the statistics

$$\langle \eta(t) \rangle = 0, \quad \langle \eta(t)\eta(t') \rangle = 2vE_0^2 \delta(t - t'). \quad (74)$$

Formally solving eq. (73),

$$E^*(t + \tau) = E^*(t)e^{-\tau/\tau_{\text{sat}}} + \mathcal{E}(1 - e^{-\tau/\tau_{\text{sat}}}) + \int_{t_1}^{t_2} dt' e^{-(\tau-t')/\tau_{\text{sat}}} \eta(t'), \quad (75)$$

and evaluating the noise correlations (74), we obtain the average peak value with an initial condition  $E^*(t) = E_a$  and the autocorrelation function of the fitness peak in the stationary ensemble,

$$\langle E^*(t + \tau) \rangle(E_a) = E_a e^{-\tau/\tau_{\text{sat}}} + \mathcal{E}(1 - e^{-\tau/\tau_{\text{sat}}}), \quad (76)$$

$$\langle E^*(t)E^*(t + \tau) \rangle = \mathcal{E}^2 + E_0^2 r^2 e^{-\tau/\tau_{\text{sat}}}. \quad (77)$$

It is straightforward to check that eqs. (76) and (77) are valid also for punctuated seascapes.

**Moments of the trait mean.** The Langevin equation for  $\Gamma(t)$  reads

$$\partial_t \Gamma(t) = -2\mu(\Gamma(t) - \Gamma_0) - \langle \Delta \rangle \frac{2c}{E_0^2} (\Gamma(t) - E^*(t)) + \xi(t), \quad (78)$$

where  $\xi(t)$  is a Gaussian noise with the statistics

$$\langle \xi(t) \rangle = 0, \quad \langle \xi(t)\xi(t') \rangle = \frac{\langle \Delta \rangle}{N} \delta(t - t'), \quad \langle \xi(t)E^*(t') \rangle = 0. \quad (79)$$

For diffusive seascapes, the last term in 79 is equivalent to  $\langle \xi(t)\eta(t') \rangle = 0$ , which implies that genetic drift and fitness seascape fluctuations are independent. The formal solution of eq. (78) reads

$$\begin{aligned} \Gamma(t + \tau) &= e^{-\tau/\tau_{\text{eq}}} \Gamma(t) + (1 - w(c))(1 - e^{-\tau/\tau_{\text{eq}}}) \Gamma_0 \\ &\quad + \int_t^{t+\tau} dt' (E^*(t')c\langle \delta \rangle + \xi(t')) e^{-(t+\tau-t')/\tau_{\text{eq}}}, \end{aligned} \quad (80)$$

where  $w(c) = [1 + 2\theta/(c\langle \delta \rangle)]^{-1}$ . In the case of a diffusive fitness seascape, we can insert the trajectory of the fitness peak  $E^*(t)$  given by eq. (75),

$$\begin{aligned} \Gamma(t + \tau) &= \Gamma(t)e^{-\tau/\tau_{\text{eq}}} + E^*(t)w(c, -v, r^2)(e^{-\tau/\tau_{\text{sat}}} - e^{-\tau/\tau_{\text{eq}}}) + \Gamma_0(1 - w(c))(1 - e^{-\tau/\tau_{\text{eq}}}) \\ &\quad + \mathcal{E}w(c, -v, r^2) \left[ (1 - e^{-\tau/\tau_{\text{sat}}}) + \frac{\tau_{\text{eq}}}{\tau_{\text{sat}}} (1 - e^{-\tau/\tau_{\text{eq}}}) \right] \\ &\quad + \int_t^{t+\tau} dt' \left[ \xi(t')e^{-(t+\tau-t')/\tau_{\text{eq}}} + \eta(t')w(c, -v, r^2)(e^{-(t+\tau-t')/\tau_{\text{sat}}} - e^{-(t+\tau-t')/\tau_{\text{eq}}}) \right]; \end{aligned} \quad (81)$$

see, e.g., section 4 of [52]. Evaluating the noise correlations (79), we obtain

$$\langle \Gamma \rangle = w(c)\mathcal{E} + (1 - w(c))\Gamma_0, \quad (82)$$

$$\langle \Gamma(t)\Gamma(t+\tau) \rangle = \langle \Gamma \rangle^2 + \langle \hat{\Gamma}^2 \rangle e^{-\tau/\tau_{\text{eq}}} + r^2 E_0^2 w(c, v, r^2) w(c, -v, r^2) (e^{-\tau/\tau_{\text{sat}}} - e^{-\tau/\tau_{\text{eq}}}), \quad (83)$$

$$\begin{aligned} \langle \Gamma(t+\tau)E^*(t) \rangle &= \langle \Gamma \rangle \mathcal{E} + E_0^2 r^2 w(c, -v, r^2) e^{-\tau/\tau_{\text{sat}}} \\ &\quad - \theta(\tau) E_0^2 v \tau_{\text{eq}}(c) w(c, v, r^2) w(c, -v, r^2) (e^{-\tau/\tau_{\text{eq}}} - e^{-\tau/\tau_{\text{sat}}}), \end{aligned} \quad (84)$$

where  $w(c, v, r^2) = [1 + (2\theta + 2N\tau_{\text{stat}}^{-1}(v, r^2))/(c\langle \delta \rangle)]^{-1}$  and  $\theta(\tau)$  is the unit step function; i.e.,  $\theta(\tau) = 1$  for  $\tau > 0$  and  $\theta(\tau) = 0$  otherwise. The relations (82) – (84) are also valid for punctuated seascapes, as can be shown by evaluating eq. (80) with the noise terms (76), (77), and (79). The time-reflection asymmetry of the cross-correlation (84) reflects the causal relation between  $\Gamma$  and  $E^*$ . The equal-time correlations reproduce the moments (30) obtained from the solution of the Fokker-Planck equation.

From the autocorrelation function (83), we immediately obtain the scaled divergence  $\langle d^{(1)} \rangle$  reported in eq. (42). For the divergence between descendent populations,  $\langle d^{(2)} \rangle$ , we additionally use the fact that the fitness fluctuations in the different lineages are independent of each other. In a diffusive fitness seascape, we have

$$\langle \eta_i(t)\eta_j(t') \rangle = \delta_{i,j} \delta(t-t') 2vE_0^2, \quad i, j = 1, 2, \quad (85)$$

which implies

$$\langle \langle E_1^*(t+\tau_1) - \langle E_1^*(t+\tau_1) \rangle \rangle \langle E_2^*(t+\tau_2) - \langle E_2^*(t+\tau_2) \rangle \rangle \rangle = 0; \quad (86)$$

the latter relation is valid also for punctuated seascapes.

**Propagators.** We recall the decomposition of the bivariate propagator,

$$G_\tau(\Gamma, E^* | \Gamma_a, E_a^*) = G_\tau(\Gamma | \Gamma_a, E_a^*, E^*) G_\tau(E^* | E_a^*), \quad (87)$$

which reflects the independence of the fitness peak dynamics from the trait mean.

The fitness peak propagator takes the standard form for an Ornstein-Uhlenbeck process and a Poisson jump process, respectively,

$$G_\tau(E^* | E_a^*) = \begin{cases} \frac{1}{\sqrt{2\pi\langle \hat{E}^{*2} \rangle(\tau, E_a^*)}} \exp \left[ -\frac{(E^* - \langle E^* \rangle(\tau, E_a^*))^2}{2\langle \hat{E}^{*2} \rangle(\tau)} \right] & \text{(diffusive seascape)} \\ e^{-\tau/\tau_{\text{sat}}} \delta(E^* - E_a^*) + (1 - e^{-\tau/\tau_{\text{sat}}}) R_{\text{eq}}(E^*) & \text{(punctuated seascape);} \end{cases} \quad (88)$$

see, e.g., ref. [52]. In both cases, the propagator has the same mean and variance,

$$\langle E^* \rangle(\tau, E_a^*) = E_a e^{-\tau/\tau_{\text{sat}}} + \mathcal{E}(1 - e^{-\tau/\tau_{\text{sat}}}), \quad \langle \hat{E}^{*2} \rangle(\tau) = r^2 E_0^2 (1 - e^{-\tau/\tau_{\text{sat}}}), \quad (89)$$

in accordance with eqs. (76) and (77).

For diffusive seascapes, we can also compute the Gaussian propagator of the trait mean for given fitness peak positions,

$$G_\tau(\Gamma | \Gamma_a, E_a^*, E^*) = \frac{1}{\sqrt{2\pi\langle \hat{\Gamma}^2 \rangle(\tau)}} \exp \left[ -\frac{1}{2} \frac{(\Gamma - \langle \Gamma \rangle(\Gamma_a, E_a^*, E^*, \tau))^2}{\langle \hat{\Gamma}^2 \rangle(\tau)} \right]. \quad (90)$$

If  $\tau_{\text{eq}} \lesssim \tau_{\text{sat}}(v, r^2)$  and  $\tau \lesssim \tau_{\text{sat}}(v, r^2)$ , we can approximate the stochastic trajectory of the trait optimum  $E^*(t')$  in the time interval  $t_a = t - \tau \leq t' \leq t$  by the most likely trajectory for given initial and the final

values:  $E^*(t') = E_a^* + ((t' - t_a)/\tau)(E^* - E_a^*)$ . In this saddle-point approximation, we obtain the conditional trait moments

$$\begin{aligned} \langle \Gamma \rangle(\Gamma_a, E_a^*, E^*, \tau) &= \Gamma_a e^{-\tau/\tau_{\text{eq}}} + (E_a^* w(c) + \Gamma_0(1 - w(c)))(1 - e^{-\tau/\tau_{\text{eq}}}) \\ &\quad + \frac{E^* - E_a^*}{\tau} w(c) [\tau - \tau_{\text{eq}}(1 - e^{-\tau/\tau_{\text{eq}}})], \end{aligned} \quad (91)$$

$$\langle \hat{\Gamma}^2 \rangle(\tau) = E_0^2 \frac{w(c)}{2c} (1 - e^{-\tau/\tau_{\text{eq}}}). \quad (92)$$

Equations (87) – (92) determine the joint propagator  $G_\tau(\Gamma, E^* | \Gamma_a, E_a^*)$  for divergence times  $\tau_{\text{eq}} \lesssim \tau_{\text{sat}}(v, r^2)$ . In the large-time limit,  $\tau \gg \tau_{\text{sat}}(v, r^2)$ , the propagator becomes independent of the initial condition and approaches the stationary distribution,  $G_\tau(\Gamma, E^* | \Gamma_a, E_a^*) \simeq Q_{\text{stat}}(\Gamma, E^*)$ , given by eqs. (28–30). In most biological experiments, the trait optimum values are hidden variables of the evolutionary process. In that case, the only observable propagator is the marginal propagator for the trait mean,

$$\begin{aligned} G_\tau(\Gamma | \Gamma_a) &\equiv \int dE_a^* dE^* G_\tau(\Gamma, E^* | \Gamma_a, E_a^*) \frac{Q_{\text{stat}}(\Gamma, E_a^*)}{Q_{\text{stat}}(\Gamma)} \\ &= \frac{1}{\sqrt{2\pi \langle D^{(1)} \rangle(\tau)}} \exp \left[ -\frac{1}{2} \frac{(\Gamma - \langle \Gamma \rangle(\Gamma_a, \mathcal{E}, \tau))^2}{\langle D^{(1)} \rangle(\tau)} \right]. \end{aligned} \quad (93)$$

## B Numerical simulations

We test our analytical results by simulations of a Fisher-Wright process for the evolution under neutral mutation-drift dynamics, in fitness landscapes with stabilizing selection, and in diffusive or in punctuated fitness seascapes for sexual and asexual populations. We evolve a population of  $N$  individuals with genomes  $\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(N)}$ , which are bi-allelic sequences of length  $\ell$ . A genotype  $\mathbf{a}$  defines a phenotype  $E(\mathbf{a}) = \sum_{i=1}^{\ell} E_i a_i$ ; the phenotypic effects  $E_i$  are drawn from various distributions (see below). In each generation, the sequences undergo point mutations with a probability  $\epsilon\mu$  per generation, where  $\epsilon$  is the generation time. The sequences of next generation are then obtained by multinomial sampling; the general form of the sampling probability is proportional to  $[1 + \epsilon f(E(\mathbf{a}), t)]$ , with the fitness seascape  $f(E, t)$  given by (15). For a diffusive seascape, a new optimal trait value  $E^*(t)$  is drawn before each reproduction step from a Gaussian distribution with mean  $(1 - \epsilon v/r^2)E^*(t) + \epsilon(v/r^2)\mathcal{E}$  and variance  $\epsilon v E_0^2$ . For a punctuated seascape, a new, uncorrelated fitness peak is drawn from the distribution  $R_{\text{eq}}(E^*)$  with probability  $\epsilon v/r^2$ .

The evolutionary statistics of the trait mean depends weakly but systematically on the recombination rate; this dependence arises because the mean diversity  $\langle \Delta \rangle$  enters the quasi-neutral dynamics of  $\Gamma$  [1]. To simulate evolution with a finite recombination rate  $\rho$ , we recombine the genomes of pairs of individuals with probability  $\epsilon\rho$  at a single random crossover position of the genome. For the simulation of free recombination, we randomly shuffle the alleles  $a_i^1, \dots, a_i^N$  between the individuals at each genomic site  $i$  and in each generation. Analytical and numerical results for the scaled divergence  $\langle d^{(1)} \rangle(\tau)$ , the scaled genetic load  $2NL$ , and the scaled fitness flux  $\langle 2N\phi \rangle/\mu$  under free recombination are shown in Fig. A.1; these should be compared with the corresponding results for non-recombining populations in Figs. 3, 4, and 5.

Universality is the (approximate) independence of a summary trait observable from details of the trait's genomic encoding and of its molecular evolution [25]. In Fig. A.2, we report three universality tests for the divergence-diversity ratio  $\Omega^{(1)}(\tau)$ . First, simulations show that the  $\Omega$  statistics depends only weakly on the recombination rate throughout the crossover between asexual evolution ( $\rho = 0$ ) and free recombination ( $\rho \rightarrow \infty$ ). Second, the  $\Omega$  ratio is invariant under variations in the number of constitutive genomic sites,  $\ell$ , at constant selection parameters  $c$  and  $v$ . Third, this ratio is also invariant under variations of the phenotypic effect sizes  $E_i$  at these sites; this is tested by comparing simulations for two distributions of effect sizes.

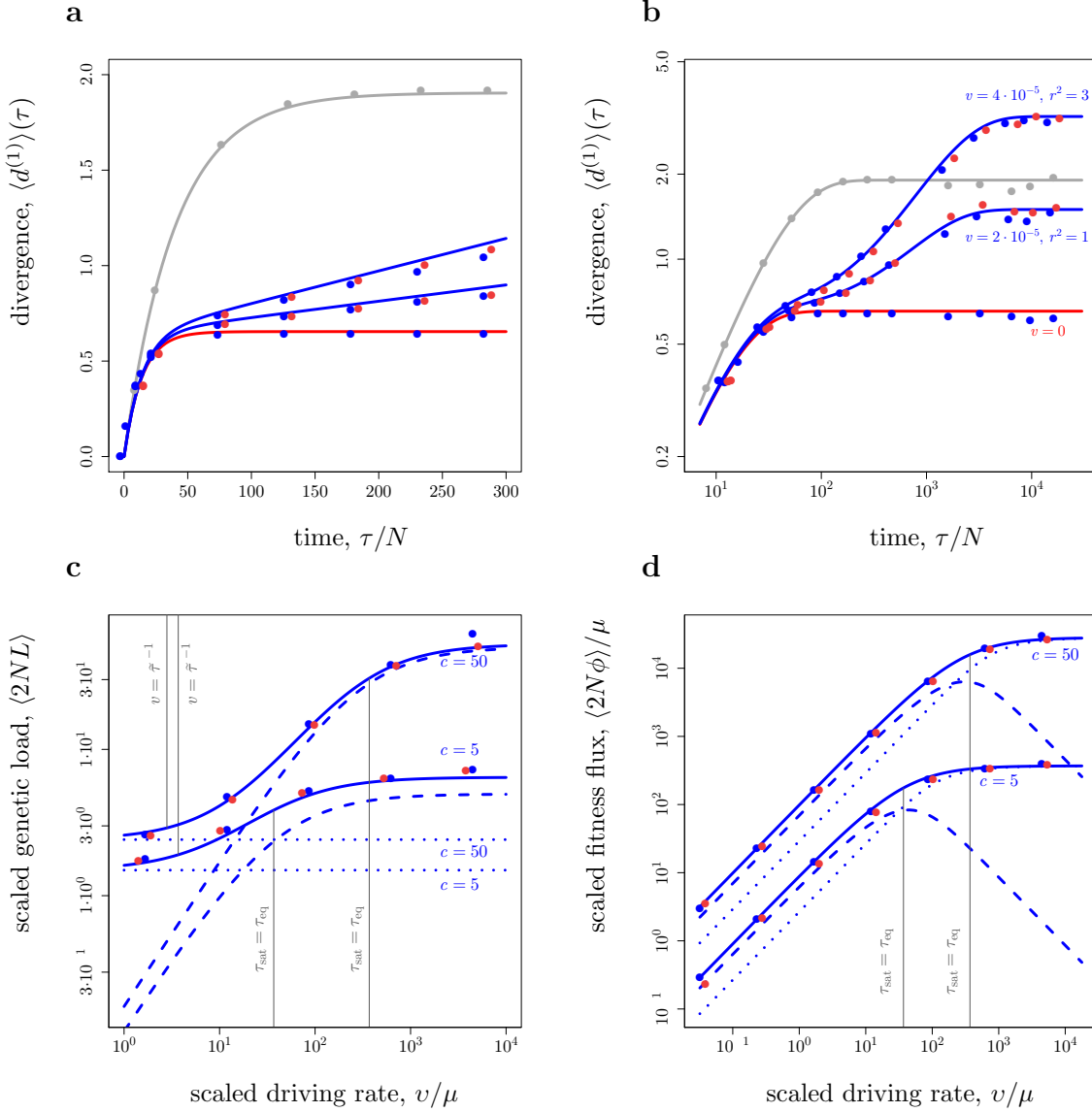


Figure A.1: **Trait evolution under free recombination.** (a,b) The scaled average divergence  $\langle d^{(1)} \rangle(\tau)$  is shown as a function of the scaled divergence time  $\tau/N$  for three cases: neutral evolution ( $c = 0$ ; grey lines), conservation in a static fitness landscape ( $c = 1, v = 0$ ; red line), and adaptation in a macro-evolutionary fitness seascape ( $c = 1, v > 0$ ; blue lines). The analytical results of eq. (42) (lines) are compared to simulation results for evolution with free recombination in diffusive and punctuated fitness seascapes (blue and red dots, respectively). The analytical value of  $\langle \delta \rangle$  is taken from eq. (69) of ref. [1]; the other parameters are as in Fig. 3. (c) Scaled genetic load  $2NL$  (full lines), adaptive load  $2NL_{\text{ad}}$  (dashed lines), and equilibrium load  $2NL_{\text{eq}}$  (dotted lines), plotted against the scaled driving rate  $v/\mu$ . The other parameters are as in Fig. 5(a). (d) Scaled fitness flux  $\langle 2N\phi \rangle$  and its components  $\langle 2N\phi_{\text{micro}} \rangle$  and  $\langle 2N\phi_{\text{macro}} \rangle$  (with decomposition constant  $k = 2$ ), plotted against the scaled driving rate  $v/\mu$ . The other parameters are as in Fig. 5(b).

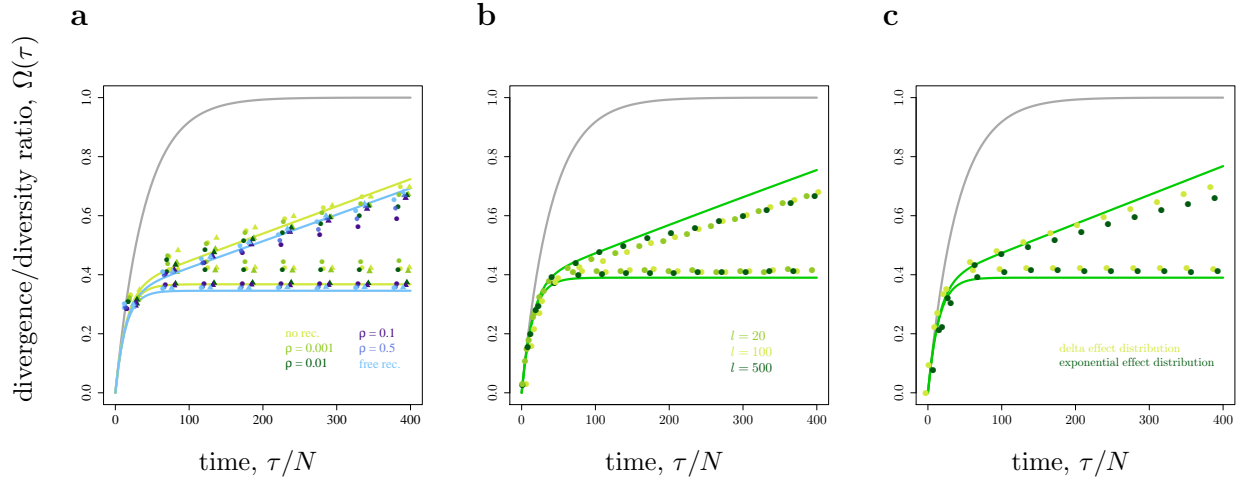


Figure A.2: **Universality of the divergence/diversity ratio  $\Omega(\tau)$ .** Numerical results for the evolution in fitness seas (  $c = 1$ ,  $v = 4 \cdot 10^{-5}$ , upper lines and dots) and fitness landscapes (  $c = 1$ ,  $v = 0$ , lower lines and dots) under different molecular conditions are compared to the analytical solutions for nonrecombining ( $\rho = 0$ ) and free-recombining ( $\rho \rightarrow \infty$ ) genomes. (a) Evolution with different recombination rates (color-coded dots for diffusive seas and triangles for punctuated seas). (b) Evolution with different numbers  $\ell$  of constitutive sites in nonrecombining populations. (c) Evolution with different effect distributions. The trait amplitudes  $E_i$  ( $i = 1, \dots, \ell$ ) are drawn from an exponential distribution with expectation value  $1/\sqrt{2}$  and from a delta distribution (all sites have amplitude  $E_i = 1$ ).